



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 171654

TO: Celine Qian
Location: REM-2A64/2C70
Art Unit: 1636
Monday, November 21, 2005

Case Serial Number: 10/099663

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

This Page Blank (uspio)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



This Page Blank (1 of 10)

A17N, Ed Hart

ACCESS DB #

FOR OFFICIAL USE ONLY

171655

PLEASE PRINT CLEARLY
RECEIVED

Scientific and Technical Information Center

NOV 16 2005

CRFG

SEARCH REQUEST FORM

CH/CHEM. DIVISION

(STIC)

Requester's Full Name: Celine Qian Examiner #: 78710 Date: 11/15/05

Art Unit: 1636 Phone Number: 2-0777 Serial Number: 101099663

Location (Bldg/Room#): 2A64 (Mailbox #): 2070 Results Format Preferred (circle): PAPER DISK

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: Conditional K10 meth. for Gene trapping & gene targeting using
Gene inducing silencer

Inventors (please provide full names):

As Kew et al.

Earliest Priority Date: 5/30/03

Search Topic:

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known.

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search SEQ ID NO: 1

1-2381 NA
LB

STAFF USE ONLY

Searcher: _____

Type of Search

1 NA Sequence (#)

Vendors and cost where applicable

_____ STN _____ Dialog

Searcher Phone #: _____

_____ AA Sequence (#)

_____ Questel/Orbit _____ Lexis/Nexis

Searcher Location: _____

_____ Structure (#)

_____ Westlaw _____ W/VV/Internet

Date Searcher Picked Up: 11/16/05

_____ Bibliographic

PS In-house sequence systems

Date Completed: 11/21/05

_____ Litigation

☒ Commercial _____ Oligomer _____ Score/Length
☒ Interference _____ SPDI _____ Encode/Transl
Other (specify) _____

Searcher Prep & Review Time: _____

_____ Fulltext

Online Time: _____

_____ Other

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 02:06:40 ; Search time 2413.28 Seconds
(without alignments)
10280.216 Million cell updates/sec

Title: US-10-099-663-1_COPY_1115_1626

Sequence: 1 ccgtgagctctctctctctc.....gtctgctctacagacagaag 512

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ha:*
2: gb_hcg:*
3: gb_hn:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the total being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	9.0	163533	9	AC069082	AC069082 Homo sapi
2	45.6	8.9	258050	1	AP003189	AP003189 Clostridi
3	45.4	8.9	1141	6	AX083744	AX083744 Sequence
4	45.4	8.9	238301	2	AC094376	AC094376 Rattus no
5	44.6	8.7	69335	2	AC026485	AC026485 Homo sapi
6	44.4	8.7	52915	2	AC127531	AC127531 Homo sapi
7	43.8	8.6	83110	2	PFMAL13PB	PFMAL13PB Homo sapi
8	43.8	8.6	110000	2	PFMAL13_13	PFMAL13_13 Plasmodu
9	43.8	8.6	147784	2	PFMAL13_14	PFMAL13_14 Continuat
10	43.6	8.6	110000	2	AL116371	AL116371 Human DNA
11	43.6	8.5	110000	3	PFMAL1P2_2	PFMAL1P2_2 Continuat
12	43.6	8.5	149266	2	CR450691	CR450691 (3 of
13	43.6	8.5	164505	2	AL330999	AL330999 Danio rer
14	43.6	8.5	164595	2	AL330999	AL330999 Homo sapi
15	43.6	8.5	175463	9	AL591686	AL591686 Homo sapi
16	43.6	8.5	186256	2	AC012079	AC012079 Homo sapi
17	43.6	8.5	266544	3	AC116956	AC116956 Dictyoste
18	43.4	8.5	174986	9	HS277662	HS277662 Homo sapi
19	43.4	8.5	198162	2	CR450817	CR450817 Danio rer

C	20	43.4	8.5	21815	2	AC074005	AC074005 Homo sapi
C	21	42.8	8.4	8868	9	AC087429	AC087429 Homo sapi
C	22	42.8	8.4	187003	2	AC023353	AC023353 Homo sapi
C	23	42.8	8.4	208232	10	AC140982	AC140982 Mus muscu
C	24	42.6	8.3	2000	6	AX655393	AX655393 Sequence
C	25	42.6	8.3	2060	8	SCU00801	SCU00801 Saccharomyc
C	26	42.6	8.3	107710	8	AC144517	AC144517 Medicago
C	27	42.6	8.3	140415	5	CR392037	CR392037 Zebrafish
C	28	42.6	8.3	199086	2	CR388073	CR388073 Danio rer
C	29	42.6	8.3	230976	2	BX927201	BX927201 Danio rer
C	30	42.6	8.3	254733	3	AC117075	AC117075 Dictyoste
C	31	42.4	8.3	5908	6	AX251915	AX251915 Sequence
C	32	42.4	8.3	5908	6	AX277931	AX277931 Sequence
C	33	42.4	8.3	5908	6	AX323618	AX323618 Sequence
C	34	42.4	8.3	110000	2	PFMAL1P1_05	PFMAL1P1_05 Continuat
C	35	42.4	8.3	149230	9	AC009947	AC009947 Homo sapi
C	36	42.4	8.3	173242	9	AC006991	AC006991 Homo sapi
C	37	42.4	8.3	175330	2	CR352298	CR352298 Homo sapi
C	38	42.4	8.3	205237	9	AC008175	AC008175 Homo sapi
C	39	42.4	8.3	252420	3	AE014841	AE014841 Plasmodu
C	40	42.2	8.2	1257	6	AR510098	AR510098 Sequence
C	41	42.2	8.2	135553	9	AC007130	AC007130 Homo sapi
C	42	42.2	8.2	150528	2	AC149242	AC149242 Macaca mu
C	43	42.2	8.2	153711	2	CR548631	CR548631 Danio rer
C	44	42.2	8.2	164439	3	AC092396	AC092396 Drosophila
C	45	42.2	8.2	176436	2	AC148535	AC148535 Macaca mu

ALIGNMENTS

RESULT 1	AC069082	163533 bp	DNA	linear	PRI 23-JAN-2002
LOCUS	AC069082	Homo sapiens chromosome 15, clone RP11-716C8, complete sequence.			
DEFINITION	AC069082				
ACCESSION	AC069082.9	GI:18182761			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens (human)				
	Homo sapiens				
REFERENCE					
AUTHORS	Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Baetien, V., Beda, F., Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deyrell, K., Dwar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPherson, R., Melatim, J., Menes, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Rhomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teste, S., Theodore, J., Tittel, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.				
TITLE					
JOURNAL					
REFERENCE					
AUTHORS	Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,				

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chopel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaidyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margitis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPherson, R., Meldrum, J., Meneus, L., Milnova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupp, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 17, 2002 this sequence version replaced gi:17059640.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L10030
Center clone name: 716_C_8

Location/Qualifiers
1..163533
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-716C8"
/clone_lib="RP11-716C8 Human Male BAC"

repeat_region
517..666
/rpt_family="TCCA)n"
complement(670..796)
/rpt_family="L2"
complement(1075..1263)
/rpt_family="MIR3"
complement(2041..2085)
/rpt_family="L2"
2219..2293
/rpt_family="MIR"
complement(2332..2411)
/rpt_family="L2"
3044..3075
/rpt_family="(GA)n"
complement(3336..3424)
/rpt_family="MIR"
complement(3426..3535)
/rpt_family="L2"
4084..4177
/rpt_family="L2"
complement(4425..4629)
/rpt_family="MIR"
4868..4911
/rpt_family="(CAGAGA)n"
5402..5578
/rpt_family="MIR3"
5580..5637

repeat_region
complement(5638..5713)
/rpt_family="L1MCA"
6194..6321
/rpt_family="L2"
7681..7843
/rpt_family="MIR"
7849..7872
/rpt_family="AT-rich"
8396..8447
/rpt_family="CT-rich"
complement(8456..8777)
/rpt_family="AluSx"
complement(9376..9682)
/rpt_family="AluY"
10920..11043
/rpt_family="MER5A"
11235..11363
/rpt_family="MER5A"
complement(12202..12268)
/rpt_family="MER5B"
12747..12861
/rpt_family="MER81"
13306..13396
/rpt_family="MER5B"
13397..13422
/rpt_family="(GGGGA)n"
13423..13488
/rpt_family="MER5B"
14336..14471
/rpt_family="MIR"
15241..15281
/rpt_family="(TG)n"
16465..16689
/rpt_family="L1MCA4"
16700..16766
/rpt_family="(TRA)n"
17990..18044
/rpt_family="L2"
18475..18600
/rpt_family="L2"
18785..18937
/rpt_family="L2"
complement(19202..19270)
/rpt_family="L2"
19524..19676
/rpt_family="MLT1J2"
complement(19677..19841)
/rpt_family="L2"
complement(20129..20206)
/rpt_family="MIR"
20775..21023
/rpt_family="MIR"
complement(22155..22575)
/rpt_family="L2"
22873..22987
/rpt_family="L2"
23520..23568
/rpt_family="MIR"
25225..25379
/rpt_family="MIR3"
complement(25438..25634)
/rpt_family="MIR"
26260..26378
/rpt_family="MIR3"
26592..26622
/rpt_family="GC-rich"
28026..28045
/rpt_family="(CA)n"
28046..28077
/rpt_family="(CA)n"
28476..28502
/rpt_family="(TTTAA)n"

unure complement (28639..28644)
/note=<30 qual SNGL region"
28734..28744
unure /note=<30 qual SNGL region"
28768..28774
unure /note=<30 qual SNGL region"
28773..28793
repeat_region /rpt_family="GC-rich"
28836..28841
unure /note=<30 qual SNGL region"
28855..28867
unure /note=<30 qual SNGL region"
28974..29317
repeat_region /rpt_family="(CA)n"

Query Match 9.0%; Score 46; DB 9; Length 163533;
Best Local Similarity 45.5%; Pred. No. 1.3;
Matches 163; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 23 TTATTATGTGAGATAGATATGATTCACCAATTGAAATGACAAAGCAGATTGTG 82
DB 48902 TTTAAAGATTGATTAATAATTTTATCTTAAATTTTAAATTTTAAATTTATG 48961
QY 83 TTTTATCTGTGGTAATACGTTTCTCCAGTTGTATAAGACCCCTCCAGATTA 142
DB 48962 GATTACAGATAGATTAATGCTGTATCCAGTGTGTTAAATTTGTTAACACATTGAA 49021
QY 143 AGTCTATGACAAAGAAATGCAATACATTCCTTACGTCATTATTTTTCATTA 202
DB 49022 AATTCACTACTTTTAAAGATTAATTTTATTTATTTATTTATTTATTTGTTGAACA 49081
QY 203 GATAGCCGTTTCTTACACACTCAATTAAGATGAACAGAAATGAGGTAGTACTG 262
DB 49082 TTTTATTTTAAATTTTCATGCGTACATAGTACATTAATTTAAAGGGGTACATGATA 49141
QY 263 TTTATAAGAGATTAATAAGATTAATCATCATCTTTGAGCAATTAAGAGGAGAGAT 322
DB 49142 TATTTGATACAGACATTAATAATTTATTTTAAAGATTAATTTTCAAAAGCTTTA 49201
QY 323 TCAGCAAAAGAGTGTCTTACAGTGGAAAACAAGTTAAACTAAAGTGAACCCCTCT 380
DB 49202 TTTGATTTTGTGATTAATAATTTGACATAGTTTAAAGATTAATGATCTCACCTCT 49259

RESULT 2
AP003189 298050 bp DNA linear BCT 20-MAR-2004
LOCUS Clostridium perfringens str. 13
DEFINITION AP003189 BAO00016
ACCESSION AP003189 GI:18146727
VERSION
KEYWORDS
SOURCE
ORGANISM
Clostridium perfringens str. 13
Clostridium perfringens str. 13
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.

REFERENCE
AUTHORS Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A.,
Shiba, T., Ogasawara, N., Hattori, M., Kuhara, S. and Hayashi, H.
TITLE Complete genome sequence of Clostridium perfringens, an anaerobic
flesh-eater
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (2), 996-1001 (2002)
MEDLINE 21664373
PUBMED 11792842
2 (bases 1 to 298050)
REFERENCE Hattori, M., Yamashita, A., Oshima, K. and Shiba, T.
AUTHORS Direct Submission
TITLE Submitted (15-PEB-2001) Masahira Hattori, Kitasato Institute for
JOURNAL Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamihara,
Kanagawa 228-8555, Japan
(E-mail: hattori@genome.1s.kitasato-u.ac.jp,
URL: http://genome.1s.kitasato-u.ac.jp, Tel: 81-42-778-8194,
Fax: 81-42-778-8193)
COMMENT On Jan 14, 2002 this sequence version replaced gi:18144662.

This work was done in collaboration with Tohru Shimizu, Kaori Ohtani, Hideo Hayaishi (University of Tsukuba), Hideki Hirakawa, Satoru Kuhara (Kyushu University), Naotake Ogasawara (Nara Institute of Science and Technology), and supported by the Research for the Future Program of the Japan Society for the Promotion of Science.

FEATURES

source
Location/Qualifiers
1..298050
/organism="Clostridium perfringens str. 13"
/mol_type="genomic DNA"
/strain="13"
/db_xref="taxon:195102"
/note="anaerobic pathogen for gas gangrene"
complement(34..1056)
/gene="CPE1002"
complement(34..1056)
/gene="CPE1002"
/note="340 aa, no significant homology."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB80708.1"
/db_xref="GI:18144663"
/translation="MSNKKINPKMNNSDKVLINSLVLDGISLKEISKIDIOKDD
IIRKMBEGPFPDSDGPKKDEKSTSDIRPSVTSKSLPKPERDVOSIISI
NEKLNKIDETTNILPEITSEKSVTKEDNEDTKTSQSTLEIPRVSRTRARK
KKKTLIKMISYFNKSFSLIKFLIKKSSSKLBCNNFKKEVNIKSTNLLIE
VAIEKLIPVNDIKETSHGNSSEKLPFEIKIPNFGRRIRIPVDDLDIEINLND
VSLDKDKDFDKNSDPIKINNSDLSNDRKSLREKVCNLEKINELEPFLIF
QAMSNK"
1361..1540
/gene="CPE1003"
1361..1540
/gene="CPE1003"
/note="59 aa, similar to gnu:AP001515.87 BH2353 gene
product from Bacillus halodurans (79 aa); 46.7% identity
in 45 aa overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB80709.1"
/db_xref="GI:18144664"
/translation="NAIDELTKRINELHKKHKEEGISEBHEKREELREYINRFSN
LKEQLKGIPEKNNK"
1552..2475
/gene="hprK"
1552..2475
/gene="hprK"
/note="307 aa, similar to gnu:AP001519.105 Hpr (Ser/Thr)
protein kinase/phosphatase from Bacillus halodurans (310
aa); 45% identity in 282 aa overlap
CPE1004"
/codon_start=1
/transl_table=11
/product="probable Hpr. Ser/Thr protein kinase/phosphatase"
/protein_id="BAB80710.1"
/db_xref="GI:18144665"
/translation="MGVTEKLIKDPSELEVIQIGBNVPIVSDVNPQLQAGFYNY
FPERIOVGKAEKSPLEMDSPDLKRLKRPSPDISCLITRGLIEBELKARK
RLMLIRSMVMTTKRISKTITWYISDMAPETRLHGVIVGIVGMLIGESGKET
ALELIRGRLVTDVAVIDKEIDGDIIGSPETITGMLFVRKGIIDVSAIGLSSTL
NSKQIKIITHFEHMKDDGDYDRLGVNDEYDILGVKKLRLVIRGRNIAVITEAA
ANVRYQMSDISPVDIIEKRLMESKEKSI"
2520..3449
/gene="CPE1005"
2520..3449
/gene="CPE1005"
/note="309 aa, similar to C-terminal of pir:S75961
magnesium/cobalt transport protein gl10507 from
Synechocystis sp. (strain PCC 6803) (387 aa); 22.4%
identity in 263 aa overlap. 2 putative transmembrane
regions were found by PSORT."
/codon_start=1

aa): 62.9% identity in 116 aa overlap
CPBI009"
/codon_start=1
/transl_table=11
/product="methylglucosyl synthase"
/protein_id="BAB80715.1"
/db_xref="GI:18144670"
/translation="MKIALYHDKKEEMIELAKDEEDKLSKHILVATGTTGKINON
TSLEVRCKSGPLGDOEIGAMVANHVDVAVIFLRPLFTQPEPDISALLRLCDVVK
VPLATINRESAKLIMADI"
7278..7613
/gene="CPBI010"
7278..7613
/note="111 aa, no significant homology. Putative
N-terminal signal sequence and 2 putative transmembrane
regions were found by PSORT."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB80716.1"
/db_xref="GI:18144671"
/translation="MRGNIQLQMSNMILVSTLITTKYILENIEICKPKKIGLTK
KTKKIVIFILTVSIVIIIGITFPYFKMFLNVAISIGNRTLIVAFQLPFGIILFLF
INRKKYKVK"
7739..7799
/gene="CPBI011"

gene

Query Match 8.9%; Score 45.6; DB 1; Length 298050;
Best Local Similarity 51.5%; Pred. No. 1.4;
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Oy 98 AAATACGTTTTCACGTGATGATTAAGACCCCTCCACAGATTAAGTCTCATGCACAA 157
Db 208707 AAGTTTTTTTATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 2087
Oy 158 AGAAATGTCATTAATCATCTCTTAAGTCTCATTAATTAATTAATTAATTAATTAAT 217
Db 208767 AAATAGATTAATAAGATTAATCTTGAACATTAATTAATTAATTAATTAATTAAT 2088
Oy 218 ACTCAACTCAATTAATTAATGAACGAATGAATGGTTAGTGCCTGTTTAAAGAAAGT 277
Db 208827 TCTTATATATGATGAGGATGATGGAATGGAACATTAATTAATTAATTAATTAAT 2088
Oy 278 AATAAGACTACTATCATCATTTGA 301
Db 208887 AATTAAACTTAATCTCTCAAGGA 208910

RESULT 3
AX083744 1141 bp DNA 11near PAT 28-FEB-2001
LOCUS AX083744
DEFINITION Sequence 22 from Patent WO011061.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Kunst, J. and Clemens, S.
TITLE Regulation of embryonic transcription in plants
JOURNAL Patent: WO 011061-A 22 15-FEB-2001;
UNIVERSITY OF BRITISH COLUMBIA (CA)
FEATURES
source
Location/Qualifiers
1..1141
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1..1141
/note="consensus sequence of A.T., L.A., and B.N. PAEL
promoters"

PROMOTER

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 237092: contig of 237092 bp in length
 * 237093 237192: gap of unknown length
 * 237193 238301: contig of 1109 bp in length.
 Location/Qualifiers
 1. 238301
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-3p1"

ORIGIN

Query Match 8.9%; Score 45.4; DB 2; Length 238301;
 Best Local Similarity 53.0%; Pred. No. 1.6; Indels 0; Gaps 0;
 Matches 97; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 173 ATTCTCTAGTCTCATTATTTATTTTATTAGATAGCGGTTTCTACTCAACTCAATA 232
 |||||
 Db 152787 ATCTTCTTTGTGTAATTAATGTAATAATTTCCCTTGATTATACATGATTTTATCT 152728
 |||||

QY 233 AGATGACAGATGATGATGGTTTGACGCTGTTTAAAGAGATTAATAAGTACTATTC 292
 |||||
 Db 152727 GAGTCGACATTTAAATGTTAGTATGCTCCCAAAAAAATATATATAAATAAT 152668
 |||||

QY 293 ATCATTTAGGCAATAGGAGGAGGAGATTTCAGCAACGTCGCTTCAAGTGGAAA 352
 |||||
 Db 152667 GGAAATAGAGCAATCAGAGAGACATGCAACACCTGGACATGAAATATCAAA 152608
 |||||

QY 353 CAA 355
 |||||
 Db 152607 GAA 152605

RESULT 5
 AC026485/c 69335 bp DNA linear HTG 22-MAR-2000
 LOCUS AC026485 Homo sapiens clone RP11-24F22, LOW-PASS SEQUENCE SAMPLING.
 DEFINITION AC026485
 ACCESSION AC026485.1 GI:7280315
 VERSION HTG; HTGS_PHASE0.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 69335)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 AUTHORS Birren, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barina, N., Bastien, V., Bedalov, F., Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G., Campoliano, A., Castelle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Coyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hago, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pletzer, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,

TITLE
JOURNAL
COMMENT

Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submission@genome.wi.mit.edu
 ----- Project Information
 Center project name: L446
 Center clone name: 24_F_22

* NOTE: This record contains 88 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 709: contig of 709 bp in length
 * 1 809: gap of 100 bp
 * 810 1511: contig of 702 bp in length
 * 1512 1611: gap of 100 bp in length
 * 1612 2309: contig of 638 bp in length
 * 2310 2409: gap of 100 bp
 * 2410 3108: contig of 639 bp in length
 * 3109 3209: gap of 100 bp
 * 3209 3893: contig of 685 bp in length
 * 3894 3993: gap of 100 bp
 * 3993 4686: contig of 693 bp in length
 * 4687 4786: gap of 100 bp
 * 4787 5482: contig of 636 bp in length
 * 5483 5582: gap of 100 bp
 * 5583 6292: contig of 710 bp in length
 * 6293 6392: gap of 100 bp
 * 6393 7092: contig of 700 bp in length
 * 7093 7192: gap of 100 bp
 * 7193 7865: contig of 673 bp in length
 * 7866 7965: gap of 100 bp
 * 7966 8657: contig of 692 bp in length
 * 8658 8757: gap of 100 bp
 * 8758 9444: contig of 687 bp in length
 * 9445 9544: gap of 100 bp
 * 9545 10237: contig of 693 bp in length
 * 10238 10337: gap of 100 bp
 * 10338 11014: contig of 677 bp in length
 * 11015 11114: gap of 100 bp
 * 11115 11778: contig of 664 bp in length
 * 11779 11878: gap of 100 bp
 * 11879 12555: contig of 677 bp in length
 * 12556 12655: gap of 100 bp
 * 12656 13360: contig of 705 bp in length
 * 13361 13460: gap of 100 bp
 * 13461 14163: contig of 703 bp in length
 * 14164 14263: gap of 100 bp
 * 14264 14934: contig of 671 bp in length
 * 14935 15034: gap of 100 bp
 * 15035 15720: contig of 686 bp in length
 * 15721 15820: gap of 100 bp
 * 15821 16514: contig of 694 bp in length
 * 16515 17304: gap of 100 bp
 * 17304: contig of 690 bp in length

17305 17404: gap of 100 bp
* 17405 18093: contig of 689 bp in length
* 18094 18193: gap of 100 bp
* 18194 18891: contig of 698 bp in length
* 18892 18991: gap of 100 bp
* 18992 19672: contig of 681 bp in length
* 19673 19772: gap of 100 bp
* 19773 20455: contig of 683 bp in length
* 20456 20555: gap of 100 bp
* 20556 21239: contig of 684 bp in length
* 21240 22026: contig of 687 bp in length
* 21340 22027: gap of 100 bp
* 22027 22127: contig of 700 bp in length
* 22127 22825: contig of 699 bp in length
* 22827 22927: gap of 100 bp
* 22927 23625: contig of 699 bp in length
* 23626 23725: gap of 100 bp
* 23726 24393: contig of 668 bp in length
* 24394 24493: gap of 100 bp
* 24494 25181: contig of 688 bp in length
* 25182 25281: gap of 100 bp
* 25282 25984: contig of 703 bp in length
* 25985 26084: gap of 100 bp
* 26085 26765: contig of 681 bp in length
* 26766 26865: gap of 100 bp
* 26866 27573: contig of 708 bp in length
* 27574 27673: gap of 100 bp
* 27674 28353: contig of 680 bp in length
* 28354 28453: gap of 100 bp
* 28454 29142: contig of 689 bp in length
* 29143 29242: gap of 100 bp
* 29243 29927: contig of 685 bp in length
* 29928 30027: gap of 100 bp
* 30028 30719: contig of 692 bp in length
* 30720 30819: gap of 100 bp
* 30820 31514: contig of 695 bp in length
* 31515 32308: contig of 694 bp in length
* 32309 32408: gap of 100 bp
* 32409 33094: contig of 686 bp in length
* 33095 33194: gap of 100 bp
* 33195 33895: contig of 701 bp in length
* 33896 33995: gap of 100 bp
* 33996 34705: contig of 710 bp in length
* 34706 34805: gap of 100 bp
* 34806 35504: contig of 699 bp in length
* 35505 35604: gap of 100 bp
* 35605 36291: contig of 687 bp in length
* 36292 36391: gap of 100 bp
* 36392 37090: contig of 699 bp in length
* 37091 37190: gap of 100 bp
* 37191 37881: contig of 691 bp in length
* 37882 37981: gap of 100 bp
* 37983 38671: contig of 690 bp in length
* 38672 38771: gap of 100 bp
* 38772 39465: contig of 694 bp in length
* 39466 39565: gap of 100 bp
* 39566 40255: contig of 690 bp in length
* 40256 40355: gap of 100 bp
* 40356 41052: contig of 697 bp in length
* 41053 41152: gap of 100 bp
* 41153 41847: contig of 695 bp in length
* 41848 41947: gap of 100 bp
* 41949 42630: contig of 683 bp in length
* 42631 42730: gap of 100 bp
* 42731 43417: contig of 687 bp in length
* 43418 43517: gap of 100 bp
* 43518 44185: contig of 668 bp in length
* 44186 44285: gap of 100 bp
* 44286 44986: contig of 701 bp in length
* 44987 45086: gap of 100 bp
* 45087 45766: contig of 680 bp in length
* 45767 45866: gap of 100 bp

45867 46530: contig of 664 bp in length
* 46531 46630: gap of 100 bp
* 46631 47323: contig of 693 bp in length
* 47324 47423: gap of 100 bp
* 47424 48022: contig of 599 bp in length
* 48023 48122: gap of 100 bp
* 48123 48778: contig of 656 bp in length
* 48779 48878: gap of 100 bp
* 48879 49579: contig of 701 bp in length
* 49580 49679: gap of 100 bp
* 49680 50387: contig of 708 bp in length
* 50388 50487: gap of 100 bp
* 50488 51192: contig of 705 bp in length
* 51193 51282: gap of 100 bp
* 51293 51980: contig of 688 bp in length
* 51981 52080: gap of 100 bp
* 52081 52784: contig of 704 bp in length
* 52785 52884: gap of 100 bp
* 52885 53572: contig of 688 bp in length
* 53573 53672: gap of 100 bp
* 53673 54365: contig of 693 bp in length
* 54366 54465: gap of 100 bp
* 54466 55136: contig of 671 bp in length

Query Match
Best Local Similarity 49.2%; Score 44.6; DB 2; Length 69335;
Matches 145; Conservative 0; Mismatches 149; Indels 1; Gaps 1;

QY 49 CACCAATTTAGATGAGCAAGAGATTTGTTTCTGTTTGGTAAATACGTTT 108
DB 62189 CAGCGCACCGCTGCGACATGTATATGATATGATTAACCTGCACAAATGTGCATGTA 62130
QY 109 CTCACAGTTGTATTAAGACCCCTCCACAGTATTAAGCTTATGCAAAAGAAATGTCA 168
DB 62129 CCTTAAACTTAAGATATATTAATAATAATAATAATAATAATAATAATAATAATAAGAT 62070
QY 169 ATACATCTCTAGTCTATATATATTTTCATTAGTATACCGGTTTTTACTTCACTCA 228
DB 62069 TAAGATCTCTAATCT-ACCTTCAGTACTCTGGGAACTTTTCTTAATATATATGA 62011
QY 229 AATAAGATGAACAGATGATGAGTGTAGTACTGTTTATTAAGAGTAAATGAATAC 288
DB 62010 CTTTATTTTCAACAGTGTATGATATTTGCTATTTTAAGAAACAAGATTTATTTATTTG 61951
QY 289 TATCATCATTTAGGCAATTAAGGAGAGAGATTCAGCAAAACAGTGTCTTACA 343
DB 61950 AAAAACCTATGATGTGAAGAGAGCAAGATGTGAACCAAGCTTATATAA 61896

RESULT 6
AC127531
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone CTD-2362F20 map 11, LOW-PASS
SEQUENCE SAMPLING.
AC127531 52915 bp DNA linear HTG 17-JUL-2002
Homo sapiens chromosome 11, clone CTD-2362F20
AC127531 GI:21886951
VERSION
KEYWORDS
HTG; HTGS PHASE0.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 52915)
Birren, B., Nuebaum, C., and Lander, E.
Homo sapiens chromosome 11, clone CTD-2362F20
Unpublished
2 (bases 1 to 52915)
Birren, B., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Bokkhalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

TITLE
JOURNAL
COMMENT

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneu, L., Mihova, T., Mlenga, V.,
Murphy, J., Naylor, J., Nguyen, C., Nicol, R., Nordu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Young, G., Zainoun, J.,
Zemek, U., Zimmer, A. and Zody, W.
Submitted (17-JUL-2002) Whitehead Institute/MIT Center for Genome
Direct Submissions
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L27588

Center clone name: 2362_F_20

* NOTE: This record contains 64 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 689 688: contig of 688 bp in length
* 789 788: gap of 100 bp
* 1510 1509: contig of 721 bp in length
* 1610 1609: gap of 100 bp
* 2355 2354: contig of 745 bp in length
* 2455 2454: gap of 100 bp
* 3159 3158: contig of 704 bp in length
* 3259 3258: gap of 100 bp
* 3989 3988: contig of 731 bp in length
* 3990 3989: gap of 100 bp
* 4090 4089: gap of 100 bp
* 4826 4825: contig of 736 bp in length
* 4926 4925: gap of 100 bp
* 5658 5657: contig of 732 bp in length
* 5758 5757: gap of 100 bp
* 6490 6489: contig of 732 bp in length
* 7299 7298: gap of 100 bp
* 7399 7398: contig of 709 bp in length
* 8133 8132: contig of 734 bp in length
* 8233 8232: gap of 100 bp
* 8963 8962: contig of 730 bp in length
* 9063 9062: gap of 100 bp
* 9791 9790: contig of 728 bp in length
* 9891 9890: gap of 100 bp
* 10636 10625: contig of 735 bp in length
* 10726 10725: gap of 100 bp
* 11457 11456: contig of 731 bp in length
* 11557 11556: gap of 100 bp
* 12284 12283: contig of 727 bp in length
* 12884 12883: gap of 100 bp
* 13131 13130: contig of 747 bp in length
* 13231 13230: gap of 100 bp
* 13984 13983: contig of 753 bp in length
* 14084 14083: gap of 100 bp
* 14832 14831: contig of 749 bp in length
* 14833 14932: gap of 100 bp

14933 15661: contig of 729 bp in length
15762 15761: gap of 100 bp
15762 16482: contig of 721 bp in length
16483 16582: gap of 100 bp
16583 17298: contig of 716 bp in length
17299 17398: gap of 100 bp
17399 18144: contig of 746 bp in length
18145 18244: gap of 100 bp
18245 18799: contig of 735 bp in length
18799 19079: gap of 100 bp
19080 19079: contig of 718 bp in length
19080 19797: gap of 100 bp
19798 19899: gap of 100 bp
19898 20621: contig of 724 bp in length
20622 20721: gap of 100 bp
20722 21470: contig of 749 bp in length
21471 21570: gap of 100 bp
21571 22316: contig of 746 bp in length
22317 22416: gap of 100 bp
22417 23125: contig of 709 bp in length
23126 23325: gap of 100 bp
23226 23968: contig of 743 bp in length
23969 24068: gap of 100 bp
24069 24794: contig of 726 bp in length
24894 24894: gap of 100 bp
24895 25621: contig of 727 bp in length
25622 25721: gap of 100 bp
25722 26453: contig of 732 bp in length
26454 26553: gap of 100 bp
26553 27300: contig of 747 bp in length
27301 27400: gap of 100 bp
27401 28125: contig of 725 bp in length
28126 28825: gap of 100 bp
28825 29074: contig of 749 bp in length
29074 29074: gap of 100 bp
29075 29816: contig of 742 bp in length
29817 29916: gap of 100 bp
29917 30663: contig of 747 bp in length
30664 30763: gap of 100 bp
30764 31497: contig of 734 bp in length
31498 31597: gap of 100 bp
31598 32331: contig of 734 bp in length
32332 32431: gap of 100 bp
32432 33159: contig of 728 bp in length
33160 33259: gap of 100 bp
33259 33988: contig of 729 bp in length
33989 34088: gap of 100 bp
34089 34835: contig of 747 bp in length
34836 34935: gap of 100 bp
34936 35671: contig of 736 bp in length
35672 35771: gap of 100 bp
35771 36505: contig of 734 bp in length
36506 36605: gap of 100 bp
36606 37286: contig of 681 bp in length
37287 37386: gap of 100 bp
37387 38115: contig of 729 bp in length
38116 38215: gap of 100 bp
38216 38944: contig of 729 bp in length
38945 39044: gap of 100 bp
39045 39765: contig of 721 bp in length
39766 39865: gap of 100 bp
39866 40599: contig of 734 bp in length
40599 40699: gap of 100 bp
40700 41441: contig of 742 bp in length
41442 41541: gap of 100 bp
41542 42157: contig of 616 bp in length
42158 42257: gap of 100 bp
42258 42989: contig of 732 bp in length
42989 43089: gap of 100 bp
43090 43805: contig of 716 bp in length
43806 43905: gap of 100 bp
43906 44654: contig of 749 bp in length
44655 44754: gap of 100 bp
44755 45498: contig of 744 bp in length

45499: gap of 100 bp
45599: contig of 737 bp in length
46335: gap of 100 bp
46336: gap of 100 bp
46435: contig of 714 bp in length
46436: gap of 100 bp
47149: contig of 731 bp in length
47250: gap of 100 bp
47250: contig of 731 bp in length
47981: gap of 100 bp
48080: contig of 735 bp in length
48820: gap of 100 bp
48919: contig of 727 bp in length
49646: gap of 100 bp
49746: contig of 727 bp in length
50474: gap of 100 bp
50573: contig of 695 bp in length
51268: gap of 100 bp
51369: contig of 726 bp in length
52095: gap of 100 bp
52195: contig of 721 bp in length.

FEATURES

source

1.52915
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CTD-2362F20"
/clone_id="CTDI Human BAC"

ORIGIN

Query Match 8.7%; Score 44.4; DB 2; Length 52915;
Best Local Similarity 50.0%; Pred. No. 3.6; Indels 5; Gaps 1;

Matches 141; Conservative 0; Mismatches 136; Indels 5; Gaps 1;

QY 19 ATGCTTATTTATGTCAGTAAATGATTCACCAATTGAAAGCAAGAGAGATT 78
Db 39896 ATTCTCGTCGACGTGAGATGACATTTTAAAGCAAAAGTCCAGAGAGAGAT 39955
QY 79 TGTGTTTATCTGTGGTAAATACGTTTCTCCAGTTGATTAAGACCTCCACCACT 138
Db 39956 AAATTCAGTCCTATCTCAATTAATTTCTATCAAGAACTGGGAAACTTCCA----- 40010
QY 139 ATAAAGCTTATGTCAGCAAAAGAAATGCTCAATACATCTCTTATGTCATTATTTTC 198
Db 40011 AATTAAGTCTACTGAACTTAAGAAATTAATTAAGAAATTAATTAAGTGAAGTCTT 40070
QY 199 ATTAGATAGCCGGTCTTTTCTCACTCAATTAAGTGAAGCAAGATGAGTGTAGT 258
Db 40071 AAAAGATGAGATGAGATGAGAACTACCGTCCAGTCTGAGAAAGTGAAGATGACAA 40130
QY 259 ACTGTTTATTAAGAGAGATTAATTAAGTACTATCATCATTTG 300
Db 40131 ACTCTCAAACTACAGAGAGATGAACTGTGTAAACCTTGG 40172

RESULT 7

PFMA113PB/c 83110 bp DNA linear HTG 11-AUG-1999
LOCUS Plasmodium falciparum 3D7 chromosome 13, *** SEQUENCING IN PROGRESS
DEFINITION *** 9 unordered pieces.

ACCESSION AL109814.1 GI:5731891
VERSION AL109814.1
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M.
and Barrell, B.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1999) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK

COMMENT

For more information about this sequence or the Malaria Project,

see http://www.sanger.ac.uk/Projects/P_falciparum. IMPORTANT: This
sequence is unfinished and does not necessarily represent the
correct sequence. Work on the sequence is in progress and the
release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc.
Order of segments is not known; 800 n's separate segments.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 12609: contig of 12609 bp in length
* 12610 13409: gap of 800 bp
* 13410 15875: contig of 2466 bp in length
* 15876 16675: gap of 800 bp
* 16676 21699: contig of 5024 bp in length
* 21700 22499: gap of 800 bp
* 22500 25196: contig of 2697 bp in length
* 25197 25996: gap of 800 bp
* 25997 28165: contig of 2168 bp in length
* 28166 28965: gap of 800 bp
* 28966 37197: contig of 8233 bp in length
* 37198 37997: gap of 800 bp
* 37998 41776: contig of 3779 bp in length
* 41777 42577: gap of 800 bp
* 42578 47711: contig of 5135 bp in length
* 47712 48511: gap of 800 bp
* 48512 83110: contig of 34599 bp in length.

FEATURES

source

1.83110
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/db_xref="taxon:36329"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="13"

ORIGIN

Query Match 8.6%; Score 43.8; DB 2; Length 83110;
Best Local Similarity 46.3%; Pred. No. 4.6; Indels 0; Gaps 0;

Matches 144; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 4 TGGATCCTTCATCTCATGCTTATTAATGAGAGATGATTCACCAATTTAGATG 63
Db 598 TGTGTTTCTTCAACAATGTAATTAATTAATTAATTAATTAATTAATTAATTA 539
QY 64 GACAAAGCAGAGATTTGTTTATCTGTTGGTAAATACGTTTCTCCAGTTGATTA 123
Db 538 AACAAATTAATTTTCTTTTATTAAGGTAATTAATTAATTAATTAATTAATTA 479
QY 124 GACCTTCCACAGATTAAGTCTATGCAACAAAGAAATGCAATTCCTTTAGT 183
Db 478 AATCTATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 419
QY 184 CTCATTAATTAATTTCTTTAGATAGCCGGTCTTTTCTACCAATTAATTAATTA 243
Db 418 TTTTCTTATCTTAAATTTTCTGCAATTTTATTAATTAATTAATTAATTAATTA 359
QY 244 ATGAATGGTTAGTGAATGTTTAAAGAGATTAATTAATTAATTAATTAATTAATTA 303
Db 358 AAAAATGACAAAGAAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 259
QY 304 CAATTAAGGAG 314
Db 298 ATAAAGGGCG 288

RESULT 8

PFMAL13_13
WPCOMMENT
Sequence split into 28 fragments LOCUS PFMAL13 Accession AL844509
Fragment Name Begin End
PFMAL13_00 1 110000
PFMAL13_01 100001 210000
PFMAL13_02 200001 310000
PFMAL13_03 300001 410000
PFMAL13_04 400001 510000
PFMAL13_05 500001 610000
PFMAL13_06 600001 710000
PFMAL13_07 700001 810000
PFMAL13_08 800001 910000
PFMAL13_09 900001 1010000
PFMAL13_10 1000001 1100000
PFMAL13_11 1100001 1210000
PFMAL13_12 1200001 1310000
PFMAL13_13 1300001 1410000
PFMAL13_14 1400001 1510000
PFMAL13_15 1500001 1610000
PFMAL13_16 1600001 1710000
PFMAL13_17 1700001 1810000
PFMAL13_18 1800001 1910000
PFMAL13_19 1900001 2010000
PFMAL13_20 2000001 2110000
PFMAL13_21 2100001 2210000
PFMAL13_22 2200001 2310000
PFMAL13_23 2300001 2410000
PFMAL13_24 2400001 2510000
PFMAL13_25 2500001 2610000
PFMAL13_26 2600001 2710000
PFMAL13_27 2700001 2732359
Continuation (14 of 28) of PFMAL13 from base 1300001 (AL844509 Plasmidium falciiparum 3D7
Query Match 8.6%; Score 43.8; DB 2; Length 110000;
Best Local Similarity 46.3%; Pred. No. 4.3;
Matches 144; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 4 TGGATCCTTCATCTCATGCTTATTATGTGAGTAGATGATCTCACCAGTTGATGAA 63
DB 103932 TGTGTTTTCTTCAACAGATGATATATTAATTAATATTTTCTACATATTTATTTTA 103991
QY 64 GACAAAGCAGAGATTGTTGTTTATCTGTTGGTAAATAGCTTTCCAGTTGATGAA 123
DB 103992 AACAAATTAATTTTTTTTTTTTTTTTATTAAGAGTAAAAAATATATGTTTACATATA 104051
QY 124 GACCTCCCAACAGATGATGAGTAAAGTCTATGCAACAAAGAAATGCAATGATCTCTAGT 183
DB 104052 AATCTATATATGATTAATCTAATAAATCTATGACACCTTAATTTTATGATATTTTTTTT 104111
QY 184 CTCATATATTTTCTCATTTAGATAGCCGTTTTTCTACACACTCAATTAAGATGACAGA 243
DB 104112 TTTTCTTTTATCTTATTTTCTGTCATTTTTTATTAATTAACATAAAAAA 104171
QY 244 ATGAATGGGTAGTGACTGTTTATTAAGAAGATTAATAAGATCATTCATTGAGG 303
DB 104172 AAAAAATGACAAAGGAAATGAATAATTAATTAATAAAGAACGACGTTTGG 104231
QY 304 CAATTAAGGAG 314
DB 104232 ATAAAGGGCG 104242
RESULT 9
PFMAL13_14
WPCOMMENT
Sequence split into 28 fragments LOCUS PFMAL13 Accession AL844509
Fragment Name Begin End
PFMAL13_00 1 110000
PFMAL13_01 100001 210000
PFMAL13_02 200001 310000
PFMAL13_03 300001 410000
PFMAL13_04 400001 510000

PFMAL13_05 500001 610000
PFMAL13_06 600001 710000
PFMAL13_07 700001 810000
PFMAL13_08 800001 910000
PFMAL13_09 900001 1010000
PFMAL13_10 1000001 1110000
PFMAL13_11 1100001 1210000
PFMAL13_12 1200001 1310000
PFMAL13_13 1300001 1410000
PFMAL13_14 1400001 1510000
PFMAL13_15 1500001 1610000
PFMAL13_16 1600001 1710000
PFMAL13_17 1700001 1810000
PFMAL13_18 1800001 1910000
PFMAL13_19 1900001 2010000
PFMAL13_20 2000001 2110000
PFMAL13_21 2100001 2210000
PFMAL13_22 2200001 2310000
PFMAL13_23 2300001 2410000
PFMAL13_24 2400001 2510000
PFMAL13_25 2500001 2610000
PFMAL13_26 2600001 2710000
PFMAL13_27 2700001 2732359
Continuation (15 of 28) of PFMAL13 from base 1400001 (AL844509 Plasmidium falciiparum 3D7
Query Match 8.6%; Score 43.8; DB 2; Length 110000;
Best Local Similarity 46.3%; Pred. No. 4.3;
Matches 144; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 4 TGGATCCTTCATCTCATGCTTATTATGTGAGTAGATGATCTCACCAGTTGATGAA 63
DB 3932 TGTGTTTTCTTCAACAGATGATATATTAATTAATATTTTCTACATATTTATTTTA 3991
QY 64 GACAAAGCAGAGATTGTTGTTTATCTGTTGGTAAATAGCTTTCCAGTTGATGAA 123
DB 3992 AACAAATTAATTTTTTTTTTTTTTTTATTAAGAGTAAAAAATATATGTTTACATATA 4051
QY 124 GACCTCCCAACAGATGATGAGTAAAGTCTATGCAACAAAGAAATGCAATGATCTCTAGT 183
DB 4052 AATCTATATATGATTAATCTAATAAATCTATGACACCTTAATTTTATGATATTTTTTTT 4111
QY 184 CTCATATATTTTCTCATTTAGATAGCCGTTTTTCTACACACTCAATTAAGATGACAGA 243
DB 4112 TTTTCTTTTATCTTATTTTCTGTCATTTTTTATTAATTAACATAAAAAA 4171
QY 244 ATGAATGGGTAGTGACTGTTTATTAAGAAGATTAATAAGATCATTCATTGAGG 303
DB 4172 AAAAAATGACAAAGGAAATGAATAATTAATTAATAAAGAACGACGTTTGG 4231
QY 304 CAATTAAGGAG 314
DB 4232 ATAAAGGGCG 4242
RESULT 10
AL136371
LOCUS AL136371 147784 bp DNA linear PRI 30-NOV-2000
DEFINITION Human DNA sequence from clone RP11-110P20 on chromosome 1q31.1
ACCESSION AL136371
VERSION AL136371.8 GI:9187156
KEYWORDS HTG; Cpg Island.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 147784)
Chapman, J.
DIRECT SUBMISSION
TITLE Submitted (21-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 18A, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonequery@sanger.ac.uk
COMMENT On Jul 14, 2000 this sequence version replaced gi:8894185.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chrl>

RP11-110P20 is from the library RPCT-11.1 constructed at the

Roswell Park Cancer Institute by the group of Pieter de Jong. For

further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pBACe3.6

This sequence is the entire insert of clone RP11-110P20 The true

left end of clone RP11-487023 is at 59783 in this sequence.

Location/Qualifiers

FEATURES
source

```
1.147784
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="q31.1"
/clone="RP11-110P20"
/clone_1lb="RPCT-11.1"
8.539
/note="match: GSS: Em:AQ321011"
repeat_region
/note="L1MA7 repeat: matches 6014. .6255 of consensus"
202. .675
/note="match: GSS: Em:AQ406263"
752. .956
repeat_region
/note="MER58A repeat: matches 1. .213 of consensus"
1045. .1214
/note="MER58A repeat: matches 1. .175 of consensus"
1887. .1990
repeat_region
/note="L2 repeat: matches 2562. .2672 of consensus"
3724. .3763
repeat_region
/note="20 copies 2 mer ac 77% conserved"
3917. .4210
repeat_region
/note="AluY repeat: matches 1. .294 of consensus"
5674. .6134
repeat_region
/note="MIR1H repeat: matches 68. .512 of consensus"
7252. .7537
repeat_region
/note="AluX repeat: matches 10. .296 of consensus"
/note="complement(7552. .7954)"
misc_feature
/note="match: GSS: Em:AQ018563"
8079. .8200
/note="match: GSS: Em:AQ040820"
8206. .8595
misc_feature
/note="match: GSS: Em:AQ040820"
9363. .9625
repeat_region
/note="L1MC4 repeat: matches 7306. .7586 of consensus"
9652. .10604
repeat_region
/note="L1M1 repeat: matches -1390. .-390 of consensus"
10596. .10861
repeat_region
/note="L1M1 repeat: matches 4665. .4914 of consensus"
10849. .13488
repeat_region
/note="L1PA13 repeat: matches 3488. .6156 of consensus"
13497. .13917
repeat_region
/note="L1M1 repeat: matches 4906. .5255 of consensus"
13918. .14162
repeat_region
```

```
/note="AluSg repeat: matches 3. .247 of consensus"
14163. .14306
repeat_region
/note="L1M1 repeat: matches 5255. .5403 of consensus"
14303. .15870
repeat_region
/note="L1PA7 repeat: matches 4578. .6141 of consensus"
17029. .17153
repeat_region
/note="MIR repeat: matches 48. .183 of consensus"
complement(17194. .17431)
misc_feature
/note="match: GSS: Em:AQ627818"
complement(17203. .17626)
misc_feature
/note="match: GSS: Em:AQ382224"
17205. .17574
/note="match: GSS: Em:AQ344420"
17239. .17623
repeat_region
/note="MER4B repeat: matches 206. .574 of consensus"
17641. .17749
repeat_region
/note="AluSg/x repeat: matches 178. .286 of consensus"
17751. .18087
repeat_region
/note="MER4D repeat: matches 236. .584 of consensus"
18188. .18482
repeat_region
/note="AluY repeat: matches 1. .296 of consensus"
18630. .18980
repeat_region
/note="MER4EB repeat: matches 1. .236 of consensus"
19636. .19847
repeat_region
/note="MER20 repeat: matches 2. .217 of consensus"
21147. .22145
repeat_region
/note="PTRS repeat: matches 744. .2021 of consensus"
21337. .21759
misc_feature
/note="CpG island"
/evidence="not_experimental"
22150. .22247
repeat_region
/note="LTRJ30 repeat: matches 544. .642 of consensus"
22257. .22471
repeat_region
/note="PTRS repeat: matches 177. .392 of consensus"
22741. .22869
repeat_region
/note="L1P1 repeat: matches 5681. .5818 of consensus"
22841. .22895
repeat_region
/note="L1PA3 repeat: matches 6092. .6146 of consensus"
23141. .23286
repeat_region
/note="AluSg/x repeat: matches 1. .115 of consensus"
23402. .23802
misc_feature
/note="match: GSS: Em:AQ797421"
complement(23554. .24004)
misc_feature
/note="match: GSS: Em:AQ007707"
23816. .23939
repeat_region
/note="62 copies 2 mer aa 58% conserved"
24049. .24742
repeat_region
/note="L1MC4 repeat: matches 6595. .7295 of consensus"
24800. .25176
repeat_region
/note="MER92A repeat: matches 1. .359 of consensus"
24919. .25385
misc_feature
/note="match: GSS: Em:AQ675810"
25096. .25485
misc_feature
/note="match: GSS: Em:B38318"
25198. .25276
repeat_region
/note="TIGR2 repeat: matches 2628. .2708 of consensus"
25279. .25332
repeat_region
/note="L1PA16 repeat: matches 6103. .6157 of consensus"
25521. .30992
repeat_region
/note="L1PA15 repeat: matches 688. .6157 of consensus"
31192. .31791
repeat_region
/note="L1PA13 repeat: matches -651. .-50 of consensus"
31801. .32118
repeat_region
/note="TIGR2 repeat: matches 2338. .2658 of consensus"
32148. .32216
repeat_region
/note="TIGR2 repeat: matches 1. .69 of consensus"
32253. .32572
repeat_region
/note="AluY repeat: matches 1. .307 of consensus"
33626. .33978
repeat_region
/note="MER47B repeat: matches 11. .418 of consensus"
34298. .34625
repeat_region
/note="MER2 repeat: matches 1. .341 of consensus"
34950. .34985
repeat_region
```

```
misc_feature /note="18 copies 2 mer ct 80% conserved"
complement(35544..35956)
/note="match: GSS: Em:AQ184339"
repeat_region 36136..36231
/note="6 copies 16 mer 72% conserved"
repeat_region 36145..36230
/note="43 copies 2 mer ca 76% conserved"
repeat_region 37001..37267
/note="11usC repeat: matches 32..298 of consensus"
repeat_region 37814..37912
/note="12 repeat: matches 2608..2710 of consensus"
repeat_region 38718..38839
/note="11 repeat: matches 4706..4841 of consensus"
repeat_region 39077..39381
/note="11usP repeat: matches 1..313 of consensus"
repeat_region 39552..39750
/note="11M/C/D repeat: matches 5281..5479 of consensus"
repeat_region 40657..40784
/note="8 copies 16 mer 68% conserved"
repeat_region 40698..40785
/note="44 copies 2 mer ag 64% conserved"
misc_feature complement(40973..41511)
/note="match: GSS: Em:B55962"
misc_feature complement(41075..41505)
/note="match: GSS: Em:AQ217602"
misc_feature complement(41254..41503)
/note="match: GSS: Em:AQ076892"
misc_feature complement(41291..41511)
/note="match: GSS: Em:AQ079585"
repeat_region 41765..41965
/note="match: GSS: matches 6075..6283 of consensus"
misc_feature 42059..42402
/note="match: GSS: Em:AQ215516"
misc_feature 42059..42255
/note="match: GSS: Em:B46341"
misc_feature 42089..42255
/note="match: GSS: Em:AQ036131"
```

Query Match Best Local Similarity 8.6%; Score 43.8; DB 9; Length 147784; Matches 133; Conservative 0; Mismatches 112; Indels 2; Gaps 2;

```
QY 39 AATAGATTCCTCAACCAATTAGATGACAAAGCAGAGATTGTTTATCTGTGGTA 98
DB 35138 ACTTCACCTTTACCAAGTGAAGATTTAAATTTGTAAGAAATATTTGCTTTTA 35197
QY 99 AATACGTTTCTCCAGTTGTAAT-AAAGCCCTCCACAGATTAAGTCTATGCAACA 157
DB 35198 AAGTTAGTTCTACTTATATGAGAGGTCATTAACAATTTAAAGTTTATGTTATA 35257
QY 158 AGAAATGTCATACATCTCTTAGTCTATTAT-TTTCATTGATAGCCGGTTTTT 216
DB 35258 AGGCTCTGAAAGAAATGCTTCGCGCCATATTTATGTTTCATTAATCTGTTAG 35317
QY 217 TACTCAACTCAATTAAGATGAACAGATGATGGTTAGTGCCTGTTATAAAGAGAG 276
DB 35318 TCTTATATCTCATGTGATGAAGACCTGAAGAGGTAATTTGATGTTGAGGAAA 35377
QY 277 TAATATA 283
DB 35378 TCTTAGA 35384
```

RESULT 11
PFMALIP2_2/c
WPCOMMENT
Sequence split into 4 fragments LOCUS PFMALIP2 Accession AL031745

```
Fragment Name Begin End
PFMALIP2_0 1 110000
PFMALIP2_1 100001 210000
PFMALIP2_2 200001 310000
PFMALIP2_3 300001 384550
Continuation (3 of 4) of PFMALIP2 from base 200001 (AL031745 Plasmidium falciptarum DNA B
```

```
Query Match Best Local Similarity 8.5%; Score 43.6; DB 3; Length 110000; Matches 103; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 17 TCATGCTTATATGAGTGAAGATGATTCACCAATTAAGATGACAAAGAGAGA 76
DB 47331 TTATCCCTTTATGGGGTTAAAAAATTACATTAATAAAAAAGAAAAAGAAA 47332
QY 77 TTTGTTTATCTGTTGGGTAATACGTTTCTCCAGTTGTAATAAGACCTCCACCA 136
DB 47331 AACCGATCTATTTTGTAGTATAAAAAATATATTAATTAATAATATTTATTA 47272
QY 137 GATATAAGTCCATGACCAAGAAAGTCAATCATCTCTGTCATATTATTT 196
DB 47271 GTAATACCTTATATAATAATATATTAATTCATTCATGACATCAATTTTCTTA 47212
QY 197 TCATTAGATAGCCGGTTTTTA 218
DB 47211 GAACAATATAGCTATATTTA 47190
```

RESULT 12
CR450691 149266 bp DNA linear HTG 24-MAY-2004
LOCUS Danio rerio clone CH211-77B10, WORKING DRAFT SEQUENCE, 6 unordered
DEFINITION
ACCESSION CR450691 GI:47604349
VERSION CR450691.1
KEYWORDS HTG; HTGS; PHASE1; HTGS DRAFT; HTGS FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 149266)
AUTHORS Burton, J.
TITLE Submitted (21-MAY-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
JOURNAL Genome Center

COMMENT

Center: Wellcome Trust Sanger Institute
Center site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
Project name: Project Information
Center project name: zc77B10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: dye-terminator; 100% of reads
Consensus quality: 146835 bases at least Q40
Consensus quality: 147357 bases at least Q30
Consensus quality: 147782 bases at least Q20
Insert size: 148766; sum-of-coverage
Quality coverage: 7.19% in Q20 bases; sum-of-coverage
Quality coverage: 6.80% in Q20 bases; agarose-1p

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 8085: contig of 8085 bp in length
* 8086 8185: gap of 100 bp
* 8186 48371: contig of 40186 bp in length
* 48372 48471: gap of 100 bp
* 48472 52239: contig of 3768 bp in length
* 52240 52339: gap of 100 bp

* 52340 113800: contig of 61461 bp in length
* 113801 113900: gap of 100 bp
* 113901 140425: contig of 26525 bp in length
* 140426 140525: gap of 100 bp
* 140526 149266: contig of 8741 bp in length.
Location/Qualifiers
1. 149266
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-77E10"
/clone_11b="CHORI-211"
1. 8085
/note="assembly_fragment:00038
clone_end:SP6
vector_side:left"
misc_feature
8186. 48371
/note="assembly_fragment:00595
fragment_chain:1"
48472. 52239
/note="assembly_fragment:00016
fragment_chain:1"
52340. 113800
/note="assembly_fragment:01002
fragment_chain:1"
113901. 140425
/note="assembly_fragment:00257
fragment_chain:1"
140526. 149266
/note="assembly_fragment:00146
clone_end:T7
vector_side:right"

ORIGIN

Query Match 8.5%; Score 43.6; DB 2; Length 149266;
Best Local Similarity 54.3%; Pred. No. 4.5;
Matches 88; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 62 TGGACAAAGCAGATTGTTGTTTATCGTGGGTAATACGTTTCCAGTTGTATA 121
DB 126007 TTGAAAAGCAGCAATTTTGTGTTAAATTTTAAATTAATTCACAAAGCTGATT 126066
QY 122 AGACCTCCACACATATTAAGCTCATGCAACAAGAAATGTCATCATCTCTTA 181
DB 126067 TGAATGTTTATGTTGTAATCAAAACTTACTATTAACATGATATATATGACTTG 126126
QY 182 GTCCTATTATTATTCATTAGATCCGGTCTTTTACTACA 223
DB 126127 ATTTTATTATTATTTTAAATAATTTTACATTTTTACA 126168

RESULT 13
AL390999/c 164505 bp DNA 1linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 1 clone RP11-151E7, 17 unordered pieces.
DEFINITION AL390999
ACCESSION AL390999.3 GI:10178785
VERSION HTG; HTGS_PHASE1; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Mclay, K.
TITLE Direct Submision
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Sep 16, 2000 this sequence version replaced gi:10040136.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk

Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: DA151E7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 158161 bases at least Q40
Consensus quality: 160871 bases at least Q30
Consensus quality: 161986 bases at least Q20
Insert size: 162905; sum-of-contigs
Insert size: 168883; 4.3% error; agarose-fp
Quality coverage: 4.44x in Q20 bases; sum-of-contigs Quality
coverage: 4.32x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 3194: contig of 3194 bp in length
* 3195 3294: gap of 100 bp
* 3295 7213: contig of 3919 bp in length
* 7214 7313: gap of 100 bp
* 7314 18523: contig of 11210 bp in length
* 18524 18623: gap of 100 bp
* 18624 25335: contig of 6712 bp in length
* 25336 25435: gap of 100 bp
* 25436 30801: contig of 5366 bp in length
* 30802 30901: gap of 100 bp
* 30901 39872: contig of 8971 bp in length
* 39873 39972: gap of 100 bp
* 39973 52124: contig of 12152 bp in length
* 52125 52224: gap of 100 bp
* 52225 57860: contig of 5636 bp in length
* 57861 57960: gap of 100 bp
* 57961 75786: contig of 17826 bp in length
* 75787 75887: gap of 100 bp
* 75887 100086: contig of 24200 bp in length
* 100087 100186: gap of 100 bp
* 100187 107095: contig of 6909 bp in length
* 107096 107195: gap of 100 bp
* 107196 125498: contig of 18303 bp in length
* 125499 125598: gap of 100 bp
* 125599 128839: contig of 3240 bp in length
* 128839 128938: gap of 100 bp
* 128939 147758: contig of 18820 bp in length
* 147759 147858: gap of 100 bp
* 147859 150588: contig of 2730 bp in length
* 150589 150688: gap of 100 bp
* 150689 162243: contig of 11555 bp in length
* 162244 162343: gap of 100 bp
* 162344 164505: contig of 2162 bp in length.
Location/Qualifiers
1. 164505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-151E7"
/clone_11b="RPC1-11.1"
1. 3194
/note="assembly_fragment:00973
fragment_chain:1
clone_end:SP6
vector_side:left"
3295. 7213
/note="assembly_fragment:00415
fragment_chain:1"
7314. 18523

FEATURES

source

misc_feature
1. 3194
/note="assembly_fragment:00973
fragment_chain:1
clone_end:SP6
vector_side:left"
misc_feature
3295. 7213
/note="assembly_fragment:00415
fragment_chain:1"
misc_feature
7314. 18523

```
/note="assembly_fragment:01529
fragment_chain:1"
misc_feature
/note="assembly_fragment:01161
fragment_chain:1"
misc_feature
/note="assembly_fragment:00848
fragment_chain:1"
misc_feature
/note="assembly_fragment:00318
fragment_chain:2"
misc_feature
/note="assembly_fragment:00647
fragment_chain:2"
misc_feature
/note="assembly_fragment:00648
fragment_chain:3"
misc_feature
/note="assembly_fragment:01018
fragment_chain:3"
misc_feature
/note="assembly_fragment:00762
fragment_chain:4"
misc_feature
/note="assembly_fragment:00263
fragment_chain:4"
misc_feature
/note="assembly_fragment:00161"
misc_feature
/note="assembly_fragment:00164"
misc_feature
/note="assembly_fragment:00664"
misc_feature
/note="assembly_fragment:01387"
misc_feature
/note="assembly_fragment:01428"
misc_feature
/note="assembly_fragment:01491
clone_end:T7
vector_side:right"
```

ORIGIN

```
Query Match      8.5%; Score 43.6; DB 2; Length 164505;
Best Local Similarity 55.2%; Pred. No. 4.5;
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY      138 TATTAAGTCTATGCAACAAGAAATGTCATATCTCTTATGCTCATTTATTTT 197
      |||||
Db      125236 TATTAAGCACTACCAACATGGCTACTATTTATTTATTTATTTATTTATTTAT 125177

QY      198 CATTAGATAGCCGGTTTTTACTACAACTCAATTAAGATGAAAGATGGCTTAGT 257
      |||||
Db      125176 TATTTATTTACTATTTATTTATTTCTCTCACTCTCAAAAGGCTACCAAGAAAGCATGTGACTAT 125117

QY      258 GACTGTTTATTAAGAAGATTAATTAAGATTAAT 291
      |||||
Db      125116 GACTTTATAGTACGAATTAATGTAATTAATTT 125083

RESULT 14
AC024501/c 164595 bp DNA linear HTG 04-MAY-2001
LOCUS Homo sapiens chromosome 1, clone RP11-151E7 map 1, WORKING DRAFT
DEFINITION
SEQUENCE 14 unordered pieces.
AC024501
AC024501.2 GI:9994152
VERSION
HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 164595)
Birren,B., Linton,L., Nuebaum,C. and Lander,B.
```

```
TITLE
JOURNAL
REFERENCE
AUTHORS

Homo sapiens chromosome 1, clone RP11-151E7
Unpublished
2 (bases 1 to 164595)
Birren,B., Linton,L., Nuebaum,C., Lander,B., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boudgallier,B., Brown,A., Burkett,G., Campolano,A., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M.,
Fenster,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kam,L., Karates,A.,
Klein,J., Lander,B., Laroque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McCurt,A., McKernan,K., McPheters,R., Meldrum,J.,
Menues,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M.,
Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Travers,M., Trigilio,J., Vassiliou,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.

Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 8, 2000 this sequence version replaced gi:7108297.
All repeats were identified using RepeatMasker:
Smt,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 151.E.7
Center clone name: 151.E.7

----- Summary Statistics
Sequencing vector: M13; W77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156974 bases at least Q40
Consensus quality: 161134 bases at least Q20
Insert size: 163000; agarose-fp
Insert size: 163295; sum-of-coverage
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.7 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1757 1756: contig of 1756 bp in length
1857 1856: gap of 100 bp
35722 35722: contig of 3366 bp in length
35723 35722: gap of 100 bp
35823 38369: contig of 2547 bp in length
38370 38469: gap of 100 bp
38470 42940: contig of 4471 bp in length
42941 43040: gap of 100 bp
43041 48601: contig of 5561 bp in length
48602 48701: gap of 100 bp
48702 58127: contig of 9426 bp in length
58128 58227: gap of 100 bp
58228 68052: contig of 9825 bp in length
68053 68152: gap of 100 bp
```


* 68153 80066: contig of 11914 bp in length
* 80067 80166: gap of 100 bp
* 80167 92919: contig of 12753 bp in length
* 92920 93019: gap of 100 bp
* 93020 108058: contig of 15039 bp in length
* 108059 108158: gap of 100 bp
* 108159 125753: contig of 17595 bp in length
* 125754 125853: gap of 100 bp
* 125854 142386: contig of 16533 bp in length
* 142387 142486: gap of 100 bp
* 142487 162480: contig of 19994 bp in length
* 162481 162580: gap of 100 bp
* 162581 164595: contig of 2015 bp in length.

FEATURES

SOURCE

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-15117"
/clone_id="RP11-11 Human Male BAC"
1. 1756
/note="assembly_fragment
clone end: SP6
vector_side:left"

misc_feature
1857..35722
/note="assembly_fragment"
35823..38369
/note="assembly_fragment"
38470..42940
/note="assembly_fragment"
43041..48601
/note="assembly_fragment"
48702..58127
/note="assembly_fragment"
58228..68052
/note="assembly_fragment"
68153..80066
/note="assembly_fragment"
80167..92919
/note="assembly_fragment"
93020..108058
/note="assembly_fragment"
108159..125753
/note="assembly_fragment"
125854..142386
/note="assembly_fragment"
142487..162480
/note="assembly_fragment"
162581..164595
/note="assembly_fragment
clone end: T7
vector_side:right"

ORIGIN

Query Match 8.5%; Score 43.6; DB 2; Length 164595;
Best Local Similarity 55.2%; Pred. No. 4.5;

Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 138 TATTAAGTCTTATGCAACAAAGAAATGTCATATCATCTTCTTATGTCATTTATTTT 197
DB 119464 TATAGCCACTAGCAACATGGCTACTATTTATTTATTTATTTATTTATTTATTTAT 119405
QY 198 CATTAGATAGCGGTTTTTTTACTACACTCAATTAAGATGAAGCAATGAGTGGTACT 257
DB 119404 TATTATTAATATTTATTTGCTCTCACTCTCAAAAGGCTACCGAAGCAAGTACTAT 119345
QY 258 GACTTTATTAAGAGAGTAATTAAGTACTAT 291
DB 119344 GACTTTATTAAGAGATTAATGTTAACTTAATTT 119311

RESULT 15

AL591686/c

LOCUS AL591686 175463 bp DNA linear PRI 07-JAN-2004
DEFINITION Human DNA sequence from clone RP11-561111 on chromosome 1, complete
sequence.

ACCESSION

AL591686
AL591686.10 GI:40733397

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (07-JAN-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jan 7, 2004 this sequence version replaced gi:15131352.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>
RP11-561111 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

SOURCE

1. 175463
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-561111"
/clone_id="RP11-11.2"

ORIGIN

Query Match 8.5%; Score 43.6; DB 9; Length 175463;
Best Local Similarity 55.2%; Pred. No. 4.4;

Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 138 TATTAAGTCTTATGCAACAAAGAAATGTCATATCTTCTTATGTCATTTATTTT 197
DB 99035 TATAGCCACTAGCAACATGGCTACTATTTATTTATTTATTTATTTATTTATTTAT 98976
QY 198 CATTAGATAGCGGTTTTTTTACTACACTCAATTAAGATGAAGCAATGAGTGGTACT 257
DB 98975 TATTATTAATATTTATTTGCTCTCACTCTCAAAAGGCTACCGAAGCAAGTACTAT 98916

Qy 258 GACTGTTTATAAAGAGATAATPAAGATACAT 291
| | | | | | | | | | | | | | | | | |
Db 98915 GACTTTATAGTAGGAAATAATGGTAAATCTAATTT 98882

Search completed: November 21, 2005, 07:00:36
Job time : 2422.28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 04:21:40 ; Search time 2032.64 Seconds

6292.105 Million cell updates/sec

Title: US-10-099-663-2

Sequence: 1 attattatttcattagata.....agacagaaagatgycattta 336

Scoring table: IDENTITY_NUC

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : EST: *

```
Est: *
1:  gb_est1: *
2:  gb_est2: *
3:  gb_hlc: *
4:  gb_est3: *
5:  gb_est4: *
6:  gb_est5: *
7:  gb_est6: *
8:  gb_gsb1: *
9:  gb_gsb2: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	41.8	12.4	500	1	AU088319	AU088319 AU088319
2	41.4	12.3	444	5	BM139785	BM139785 BM139785
3	41.4	12.3	504	5	BM269615	BM269615 BM269615
4	41.4	12.3	574	1	AV680413	AV680413 AV680413
5	41.4	12.3	720	5	BM135968	BM135968 BM135968
6	41.4	12.3	725	5	BM186283	BM186283 BM186283
7	41.4	12.3	730	5	BM140184	BM140184 BM140184
8	41.4	12.3	732	5	BM127139	BM127139 BM127139
9	41.4	12.3	759	1	AV877102	AV877102 AV877102
10	41.4	12.3	767	5	BM082303	BM082303 BM082303
11	41.4	12.3	774	5	BM141206	BM141206 BM141206
12	41.1	12.2	598	7	CP250818	CP250818 esa017_e
13	40.6	12.1	594	8	AZ305625	AZ305625 100006F22
14	40.6	12.1	732	9	AG596239	AG596239 Mus musc
15	40.6	12.1	781	9	AG470649	AG470649 Mus musc
16	40.4	12.0	554	5	BM009156	BM009156 BM009156
17	40.2	12.0	522	8	AZ431375	AZ431375 1M0216U11
18	40	11.9	879	8	BH137539	BH137539 ENTNA83T7T
19	40	11.9	943	8	BH147340	BH147340 ERTQW344T
20	39.8	11.8	472	5	BM093545	BM093545 BM093545
21	39.8	11.8	507	5	BM234156	BM234156 BM234156
22	39.8	11.8	688	5	BM133492	BM133492 BM133492
23	39.8	11.8	760	5	BM399296	BM399296 BM399296
24	39.2	11.7	1101	9	CNS00E06	AL065511 Drosophila

C	25	38.8	11.5	711	9	CE141435	L1gr-g8S-
C	25	38.8	11.5	649	8	BW403626	BW403626
C	27	38.8	11.5	890	8	BH135160	BH135160
C	28	38.6	11.5	463	5	BM124142	BM124142
C	29	38.6	11.5	630	5	BM254209	BM254209
C	30	38.6	11.5	715	8	BZ393910	BZ393910
C	31	38.6	11.5	932	8	BH146427	ENTPG68TR
C	32	38.6	11.5	987	9	CNS01400	ENTPG92TR
C	33	38.4	11.4	430	2	BB784350	BB784350
C	34	38.4	11.4	615	8	BH973910	odds2e01.
C	35	38.4	11.4	755	8	BZ011773	oeh48d03.
C	36	38.4	11.4	848	9	CR114143	CR114143
C	37	38	11.2	590	7	CK383133	Reverse 8
C	38	37.8	11.2	391	8	AK2002591	1a117g08.
C	39	37.8	11.2	478	2	BP286550	RPci -23-2
C	40	37.8	11.2	905	8	AZ549279	ESTF451141
C	41	37.8	11.2	928	9	CNS01761	ENTEG46TF
C	42	37.8	11.2	1200	9	CNS01671	Drogophi1
C	43	37.6	11.2	489	9	CL891587	abf51c07.
C	44	37.6	11.2	586	9	BZ948176	BZ948176
C	45	37.6	11.2	747	6	CD825484	CH240.113
							BN25.0600

ALIGNMENTS

	RESULT 1	
LOCUS	AU088319	
DEFINITION	AU088319 Sugano Malaria cDNA library Plasmodium falciparum 3D7 cDNA	500 bp mRNA linear EST 27-JAN-2001
ACCESSION	AU088319	
VERSION	AU088319	
KEYWORDS	EST.	
SOURCE	Plasmodium falciparum 3D7	
ORGANISM	Plasmodium falciparum 3D7	
REFERENCE	Eukaryotes; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.	
AUTHORS	1 (bases 1 to 500)	
TITLE	Watanabe,J., Sasaki,M., Suzuki,Y. and Sugano,S. FNU-malaria: a database for a full-length enriched CDNA library from human malaria parasite, Plasmodium falciparum	
JOURNAL	Nucleic Acids Res. 29 (1), 70-71 (2001)	
MEDLINE	20574784	
PUBMED	1125052	
COMMENT	Contact: Junichi Watanabe Institute of Medical Science The University of Tokyo, Department of Parasitology 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Tel.: 81-3-5449-5378 Fax: 81-3-5449-5410 Email: jwatanabe@manage.iims.u-tokyo.ac.jp Suzuki,Y., Yoshimoto-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library Gene 200 (1-2), 149-156 (1997).	
FEATURES	Location/Qualifiers	
Source	1..500	
	/organism="Plasmodium falciparum 3D7"	
	/mol_type="mRNA"	
	/isolate="3D7"	
	/db_xref="taxon:36329"	
	/clone="XPf6924"	
	/dev_stage="erythrocytic stage"	
	/clone_idb="Sugano Malaria cDNA library"	
ORIGIN		
	Query Match	12.4%; Score 41.8; DB 1; Length 500;
	Best Local Similarity	50.2%; Pred. No. 0.93;
	Matches 103; Conservative	0; Mismatches 102; Indels 0; Gaps 0;
OY	14 TTGATGAGCCGGTTTTTTTACTCAACAATAAATGAAGAAGAAATGATGCGTTAGTGGA 73	
b3	232 TTAATTATGATGGAATATATGAAAATGGAATAATAATAAATATACGAAATCATCTTAA 291	

QY 74 CTGTTTATAAGAGTAATAAGTACTATCATCTTTGAGGCAATAAGGAGGAGA 133
 DB 292 ATGCTCTTTATTAAGAAATCAAAATGAGAGATTCTTCAAAAAAGTTGAATAGTA 351
 QY 134 GATTGACAAACAGTGTGCTTCAAGTGGAAAAAGTTAACTAAAGTACCCCTCC 193
 DB 352 CAAGTAAATTAATAAGCTGCAACAAAGGAAAAAAGAAATGTTATCACAAGG 411
 QY 194 TTGACAAATCAATGCCACAGTTGA 218
 DB 412 TACATTAACAAACAAAAAATTGA 436
 RESULT 2
 BM139785 444 bp mRNA linear EST 03-NOV-2002
 LOCUS BM139785 Nori Satoh unpublished cDNA library, gastrula and neurula
 DEFINITION Clona intestinalis cDNA clone rcign053020 3', mRNA sequence.
 ACCESSION BM139785
 VERSION BM139785.1 GI:24496807
 KEYWORDS EST.
 SOURCE Clona intestinalis
 ORGANISM Clona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Clona.
 REFERENCE 1 (bases 1 to 444)
 AUTHORS Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 TITLE Expressed genes in Clona intestinalis (2002c)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 FEATURES
 source Location/Qualifiers
 1..444
 /organism="Clona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcign053020"
 /tissue_type="whole body"
 /dev_stage="gastrula and neurula"
 /clone_id="Nori Satoh unpublished cDNA library, gastrula and neurula"
 ORIGIN
 Query Match 12.3%; Score 41.4; DB 5; Length 444;
 Best Local Similarity 46.4%; Pred. No. 1.2; Mismatches 156; Indels 0; Gaps 0;
 Matches 135; Conservative 0;

RESULT 3
 BM269615/c 504 bp mRNA linear EST 09-NOV-2002
 LOCUS BM269615 Nori Satoh unpublished cDNA library, gastrula and neurula
 DEFINITION Clona intestinalis cDNA clone cign053020 5', mRNA sequence.
 ACCESSION BM269615
 VERSION BM269615.1 GI:24849533
 KEYWORDS EST.
 SOURCE Clona intestinalis
 ORGANISM Clona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Clona.
 REFERENCE 1 (bases 1 to 504)
 AUTHORS Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 TITLE Expressed genes in Clona intestinalis (2002c)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 FEATURES
 source Location/Qualifiers
 1..504
 /organism="Clona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cign053020"
 /tissue_type="whole body"
 /dev_stage="gastrula and neurula"
 /clone_id="Nori Satoh unpublished cDNA library, gastrula and neurula"
 ORIGIN
 Query Match 12.3%; Score 41.4; DB 5; Length 504;
 Best Local Similarity 46.4%; Pred. No. 1.2; Mismatches 156; Indels 0; Gaps 0;
 Matches 135; Conservative 0;

QY 35 ACAACTCAATTAAGATGAACAGATGAATGGTTAGTGACTGTTTATAAGAGAGTAAT 94
 DB 414 ATAACTGATTCACGAAAGCAGAACATATGTGTCCAAACGGTCAACAAATTAAGTAAT 355
 QY 95 AAAGTACTATCATCTTTGAGGCAATAAGGAGGAGAGATTTCAGCAACAGTGTCTT 154
 DB 354 GACAGTCCAAATACCTTTCTTGTAAATAATGAACCTGCTTCAATTTACATTGCACCA 295
 QY 155 ACAAGTGAAGAAACAAGTTAACTAAAGTGACCCCTCCTGACAAAGATCAATGCCACAG 214
 DB 294 CGAAATGGAATAGGTTTAAATAGTAACATTAATAATTTGTTACCGAAATGCACCCCTG 235
 QY 215 TTGAGCTTTAGCCACCATCATCATGTAATTTGCTTCTGATTAAGCTGTTCAATAA 274
 DB 234 CTAAAGCATCTGCTATTTGCTGATTTAATTTGCTGTTCTTGACACATAGTTGCAAC 175
 QY 275 TTCTCTTTGCAAGCTGTGCTACTTACCAAGAGTCTGCTTACAGACAGAAA 325
 DB 174 AGCTTTCTGTAATCTCCACCATGTCATTCAGTTCTGCTTGCCTCTTA 124
 RESULT 4
 AV680413 574 bp mRNA linear EST 05-OCT-2000
 LOCUS AV680413 Nori Satoh unpublished cDNA library Clona intestinalis
 DEFINITION cDNA clone rc1b12j8 3', mRNA sequence.
 ACCESSION AV680413
 VERSION AV680413.1 GI:10118412
 KEYWORDS EST.
 SOURCE Clona intestinalis
 ORGANISM Clona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Clona.

REFERENCE 1 (bases 1 to 574)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T
TITLE Expressed genes in Clona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh

FEATURES	Location/Qualifiers
source	1. .574

ORIGIN

```

RESULT 5
LOCUS      BM135968
DEFINITION BM135968 Nori Satoh unpublished cDNA library, gastrula and neurula
ACCESSION  BM135968
VERSION     BM135968
KEYWORDS    BM135968.1  GI:24492367
EST.
Ciona intestinalis
Ciona intestinalis
Eukaryote; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cloniidae; Ciona.
1 (bases 1 to 720)
Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..720
/organism="Ciona intestinalis"

```

```

/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rc19n042b07"
/tissue_type="whole body"
/dev_stage="gastrula and neurula"
/clone_lib="Nori Satoh unpublished cDNA library, gastrula
and neurula"

```

Query Match	12.3%	Score 41.4	DB 5	Length 720
Best Local Similarity	46.4%	Pred. No. 1.3		
Matches 135, Conservative	0	Mismatches 156	Indels 0	Gaps 0

BW186283/c			
LOCUS	BW186283	725 bp	mRNA linear EST 05-NOV-2002
DEFINITION	BW186283 Nori Satoh unpublished cDNA library, heart Clona intestinalis cDNA clone rcintc035k14 3', mRNA sequence.		
ACCION			

[illegible]

Db 304 ATAACTGATTCACGAAGACAGATATGTGTCACAAAGGTCAACAAATAGTAAT 245
 QY 95 AAAGATCATATCATTTGAGGCATTAAGGAGGAGATTCGCAACAGTGTGCTT 154
 Db 244 GACAGTTCATACCTTTCTGTAAATAATATGACCTGCTTATTTACATGACCA 185
 QY 155 ACAAGTGAACAACTAACTAAAGTGACCCCTCTGTCACAAATCAATGACACAG 214
 Db 184 CGAAATGGAATAGGTGTTAAATAGTAACATTTAAATTTGTTACGGAATGACACCCCTG 125
 QY 215 TTGAGCTTTAGCCAGCCACATCATCATGTAATTTGCTTTCTGTATTAAGCTGTTCAATA 274
 Db 124 CTAAAGCATCTGCTATTTGCTGTATTAATTTCTGTGTTCTTGTACAAATAGTTGCAAC 65
 QY 275 TTCTCTTGCAAGCTCTGCTACTTCCAGAACTGCTGCTACAGACAGAA 325
 Db 64 AGCTTTCTGTAAATCTCCACACCTGCTCATTTGCTGCTGCTCTTA 14

RESULT 7
 LOCUS BM140184 730 bp mRNA linear EST 03-NOV-2002
 DEFINITION BM140184 Nori Satoh unpublished cDNA library, gastrula and neurula
 Ciona intestinalis cDNA clone rcign05520 3', mRNA sequence.
 ACCESSION BM140184
 VERSION BM140184.1 GI:24497206
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.
 1 (bases 1 to 730)
 Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished (2002)
 JOURNAL
 COMMENT
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1..730
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcign05520"
 /issue_type="whole body"
 /dev_stage="gastrula and neurula"
 /clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"

FEATURES
 source

ORIGIN

Query Match 12.3%; Score 41.4; DB 5; Length 730;
 Best Local Similarity 46.4%; Pred. No. 1.3; Mismatches 156; Indels 0; Gaps 0;
 Matches 135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 35 ACACTCAATTAAGATGAACAGATGAGGTTAGTGACTGTTTAAGAAGACTAAT 94
 Db 40 ATGACGTATTTTCGAAAGCAGAACATATGTGTTCCAAACGGTCAACAATAAGTAAAT 99
 QY 95 AAAGATCACTATCATCTTTGAGGCAATTAAGGAGGAGATTCAGCAACAGTGTGCTT 154
 Db 100 GACAGTTCATACCTTTCTTTTAAATAATATGACCTGCTTATTTACATTTGACCA 159
 QY 155 ACAAGTGAACAAAGTTAACTAAAGTGACCCCTCTGTCACAAATCAATGACACAG 214
 Db 160 CGAAATGGAATAGGTGTTAAATAGTAACATTTAAATTTGTTACGGAATGACACCCCTG 219
 QY 215 TTGAGCTTTAGCCAGCCACATCATCATGTAATTTGCTTTCTGTATTAAGCTGTTCAATA 274
 Db 220 CTAAAGCATCTGCTATTTGCTGTATTTAAATTTCTGTGTTCTTGTACAAATAGTTGCAAC 279

QY 275 TTCTCTTGCAAGCTCTGCTACTTACCAAGAGTCTGCTACAGACAGAA 325
 Db 280 AGCTTTCTGTAAATCTCCACACCTTGCATTCAGTTCTGCTTCTGCTCTTA 330

RESULT 8
 LOCUS BM127139 732 bp mRNA linear EST 02-NOV-2002
 DEFINITION BM127139 Nori Satoh unpublished cDNA library, gastrula and neurula
 Ciona intestinalis cDNA clone rcign01316 3', mRNA sequence.
 ACCESSION BM127139
 VERSION BM127139.1 GI:24483538
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.
 1 (bases 1 to 732)
 Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 Expressed genes in Ciona intestinalis (2002c)
 Unpublished (2002)
 JOURNAL
 COMMENT
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1..732
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcign01316"
 /issue_type="whole body"
 /dev_stage="gastrula and neurula"
 /clone_lib="Nori Satoh unpublished cDNA library, gastrula and neurula"

FEATURES
 source

ORIGIN

Query Match 12.3%; Score 41.4; DB 5; Length 732;
 Best Local Similarity 46.4%; Pred. No. 1.3; Mismatches 156; Indels 0; Gaps 0;
 Matches 135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 35 ACACTCAATTAAGATGAACAGATGAGGTTAGTGACTGTTTAAGAAGACTAAT 94
 Db 20 ATAACTGATTCACCAAGACAGACATATGTGTTCCAAACGGTCAACAATAAGTAAAT 79
 QY 95 AAAGATCATATCATCTTTGAGGCAATTAAGGAGGAGATTCAGCAACAGTGTGCTT 154
 Db 80 GACAGTTCATACCTTTCTTTTAAATAATATGACCTGCTTATTTACATTTGACCA 139
 QY 155 ACAAGTGAACAAAGTTAACTAAAGTGACCCCTCTGTCACAAATCAATGACACAG 214
 Db 140 CGAAATGGAATAGGTGTTAAATAGTAACATTTAAATTTGTTACGGAATGACACCCCTG 199
 QY 215 TTGAGCTTTAGCCAGCCACATCATCATGTAATTTGCTTTCTGTATTAAGCTGTTCAATA 274
 Db 200 CTAAAGCATCTGCTATTTGCTGTATTTAAATTTCTGTGTTCTTGTACAAATAGTTGCAAC 259
 QY 275 TTCTCTTGCAAGCTCTGCTACTTACCAAGAGTCTGCTACAGACAGAA 325
 Db 260 AGCTTTCTGTAAATCTCCACACCTTGCATTCAGTTCTGCTTCTGCTCTTA 310

RESULT 9
 LOCUS AV877102 759 bp mRNA linear EST 08-NOV-2001
 DEFINITION AV877102 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
 intestinalis cDNA clone rcitb1k05 3', mRNA sequence.
 ACCESSION AV877102
 VERSION AV877102.1 GI:16864626

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Satoh N., Satou Y., Kohara Y. and Shin-I, T.	Expressed genes in ciona intestinalis	Unpublished (2000)	Contact: Nori Satoh Department of Zoology Kyoto University Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan Tel: 81-75-753-4081 Fax: 81-75-705-1113 Email: satoh@ascidiain.zool.kyoto-u.ac.jp.
FEATURES	source	1..759	location/Qualifiers	
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches 156; Indels 0; Gaps 0;		
	Matches	135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;		
ORIGIN	Query Match	12.3%; Score 41.4; DB 1; Length 759;		
	Best Local Similarity	46.4%; Pred No.1.3; Mismatches		

Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1..767
 /organism="Clona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcieg088c22"
 /tissue type="whole animal"
 /dev stage="egg"
 /clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN

Query Match 12.3%; Score 41.4; DB 5; Length 767;
 Best Local Similarity 46.4%; Pred. No. 1.3;
 Matches 135; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 35 ACAACTCAATTAAGATGAACAGAAATGAGTGGTTAGTACTGTTATTAAGAAGATTAAT 94
 DB 27 ATAACTGATATTCAGAAAGACAGAAATATGTGTTCCAAACGGTCACAAATTAAGTAAAT 86
 QY 95 AAAGATCAATCATCATTTGAGGCAATPAGGAGGAGAGATTCAGACAAAGTGTCTT 154
 DB 87 GACAGTTCATACCTCTTTCTTGAAAAAATATAGACCTGTCTTATTTTACATTCACCA 146
 QY 155 ACAAGTGAACAAAGATTAACCTAAAGTGACCCCTCTCTTGCACAAATGATCAATGCCACAG 214
 DB 147 CGAAATGGAATAGGTGTTAAATATGTAAACCTTAAATTTGTTACGAAATGACACCCCTG 206
 QY 215 TTGAGCTTTAGCCAGCCACATCATCATGTAATATGCTTTCCGTGATTAAGCCTGTTCAATA 274
 DB 207 CTAAGCATATCGCTATTTGCTGTATTTATTTCTGTGTCTTGACCAACATGTTGCAAC 266
 QY 275 TTCCTTTGCAAGCTCTGCTACTTACCAGAAAGTCTGCCCTACAGACAGAA 325
 DB 267 AGCTTCTGTAAATCTCACACCATTTGATTCAGTTGCTTGTGCTCTTAA 317

RESULT 11
 LOCUS BM141206 774 bp mRNA linear EST 03-NOV-2002
 DEFINITION BM141206 Nori Satoh unpublished cDNA library, gastrula and neurula
 Clona intestinalis cDNA clone rcign058h04 3', mRNA sequence.
 ACCESSION BM141206
 VERSION BM141206.1 GI:24498431
 KEYWORDS EST.
 SOURCE Clona intestinalis
 ORGANISM Clona intestinalis
 AUTHORS Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 TITLE Phlebobranchia; Cloniidae; Clona.
 COMMENT 1 (bases 1 to 774)
 Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
 Expressed genes in Clona intestinalis (2002c)
 Unpublished (2002)
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers
 1..774
 /organism="Clona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="rcign058h04"
 /tissue type="whole body"
 /dev stage="gastrula and neurula"
 /clone_lib="Nori Satoh unpublished cDNA library, gastrula
 and neurula"

FEATURES
 SOURCE

ORIGIN

DB	368	TAGGAGTGGGGGAAAAAGAAAGAACTTAAAAAAGAA	350
RESULT 14			
AG596239/c			
DEFINITION	AG596239	732 bp	DNA
ACCSSION	AG596239		linear
VERSION	AG596239.1		GSS 05-JUN-2004
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.		
TITLE	BAC end Sequences of Library MSMg01		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 732)		
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-NOV-2003) Maabihira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan		
	1-7-22, Suenhiro-chou, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan		
	(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,		
	Tel:81-45-503-9111, Fax:81-45-503-9170)		
	Clones are derived from the mouse BAC library MSMg01. For BAC		
	library availability, please contact Kuniya Abe (abe@rtc.riken.jp).		
	Tsukuba Institute, Bio Resource Center,		
	The Institute of Physical and Chemical Research (RIKEN) 3-1-1		
	Koyadai, Tsukuba, 305-0074 Japan		
	phone: 81-298-36-9189, fax: 81-298-36-9199		
	e-mail: abe@rtc.riken.jp		
COMMENT	PRIMERS		
	Sequencing : T7		
	LIBRARY		
	Vector : pBACe3.6		
	R.site 1 : EcoRI		
	R.site 2 : EcoRI.		
FEATURES			
Source	Location/Qualifiers		
	1..732		
	/organism="Mus musculus molossinus"		
	/mol_type="genomic DNA"		
	/sub_species="molossinus"		
	/db_xref="taxon:57486"		
	/clone="MSMg01-526L09.TU"		
	/sex="male"		
	/tissue_type="mixture of kidney and spleen"		
	/clone_lib="MSMg01 Mouse Male BAC Library"		
ORIGIN			
Query Match	12.1%	Score 40.6;	DB 9; Length 732;
Best Local Similarity	53.5%;	Pred. No. 2.2;	
Matches	85; Conservative	0; Mismatches	74; Indels
		Gaps	0;
OY	22	CCGGTTTTTCTCAACTCACTAAAGTGAACGAAGTAAGTGGTTAGTGACTTTAT	81
Db	656	CTGATTTATTTATCAAGAGAAAAAGAAAAACAAGGAAATTAAGAGAAAAAGTCAAAAAAGAA	597
OY	82	AAAGAGAGTATTAAGATTAATCATCATTTGAGGCAATTAAGGAGGAGAGATTTCAGC	141
Db	596	AAAGAGATGAAGAAAGAAAAAGAAAGATGAAGAGGAGGAGGAGGAGGAGGAGG	537
OY	142	AAACAGTGTCTTACAGTGGAAAAACAATTAACTAAA	180
Db	536	TAGGAGTGGGGGAAAAAGAAAGAACTTAAAAAAGAA	498
RESULT 15			
AG470649/c			
LOCUS	AG470649	781 bp	DNA
			linear
			GSS 04-JUN-2004

DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-361M24.TU, genomic survey sequence.
ACCESSION	AG470649
VERSION	AG470649.1 GI:48173781
KEYWORDS	GSS.
SOURCE	Mus musculus molossinus
ORGANISM	Mus musculus molossinus Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scutogomathi; Muridae; Murinae; Mus.
REFERENCE	1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y. BAC end Sequences of Library MSMg01
AUTHORS	Unpublished
TITLE	2 (bases 1 to 781)
REFERENCE	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y. Direct Submission
AUTHORS	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
JOURNAL	and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Shohiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp url:http://ngp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170) Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp
COMMENT	PRIMERS Sequencing : TU LIBRARY Vector : pBAC3.6 R.Site 1 : ECORI. R.Site 2 : EcoRI. Location/Qualifiers 1. .781 /organism="Mus musculus molossinus" /mol_type="genomic DNA" /sub_species="molossinus" /db_xref="taxon:57486" /clone="MSMg01-361M24.TU" /sex="male" /tissue type="mixture of kidney and spleen" /clone_lib="MSMg01 Mouse Male BAC Library"
FEATURES	source
ORIGIN	
Query Match	12.1%; Score 40.6; DB 9; Length 781;
Best Local Similarity	53.5%; Pred. No. 2.2;
Matches	85; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
OY	22 CCGGTTTTTAACTCAACTCAATAAATGAACAAGATGGAATGGCTTAACTGTCAGCTTTAT 81
Db	343 CTGATTATTTTCATAAAGAAAAGAAAAACAAGGAACTTAAAGGAAAGTTAGAAAAA 284
OY	82 AAAAAGAAGTAATAAGATACCTATCATCTTTTGAGCAATAAAGAGAGAGAGATTGAGC 141
Db	283 AAAAAGAAGTAAGAAAAGAAAAAAGAAAGTAAGAAAGGAGAGAGAGCGAGGAGG 224
OY	142 AAACAGTGTGCTTACAAGTGGAAAAACAAGTTAACTAAA 180
Db	223 TAGAGTGGGGGAGAAAAGAAAGATTAAAAAAAAGAA 185

Search completed: November 21, 2005, 08:26:22
Job time : 2005.64 secs

This Page Blank (uspto)

CC acid insert encoding a polypeptide, an expression vector that integrates

CC into a host cell (and comprising the isolated avian gut-specific gene
CC expression control region), expressing a heterologous polypeptide in a
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
CC molecule, and culturing the transfected cell in a medium suitable for
CC expression of a heterologous polypeptide under the control of an avian
CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression
CC control region encoded by the recombinant DNA molecule), a eukaryotic
CC cell transformed with the expression vector (or its progeny, which
CC expresses a heterologous polypeptide) and a transgenic avian having a
CC heterologous polynucleotide sequence comprising the nucleic acid insert.
CC The nucleic acids are useful for regulating heterologous nucleic acids in
CC transgenic avians, as probes in nucleic acid hybridisation assays for
CC detecting the IFABP gene expression control region, and for generating
CC transgenic birds. The present sequence is the Chicken intestinal fatty
CC acid binding protein, IFABP, gene, promoter.

XX Sequence 336 BP; 116 A; 62 C; 63 G; 95 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 336; DB 11; Length 336;

XX Best Local Similarity 100.0%; Pred. No. 2e-87; Mismatches 0; Gaps 0;

XX Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTATTTATTTTCATTAGATAGCCGGTTTTTTTACTACAACTCAATTAAGATGAACAGATG 60

DB 1 ATTATTTATTTTCATTAGATAGCCGGTTTTTTTACTACAACTCAATTAAGATGAACAGATG 60

QY 61 AATGGTTTGTGACCTGTTTATAAGAGATTAATTAAGTACTATCATTTTGGAGCA 120

DB 61 AATGGTTTGTGACCTGTTTATAAGAGATTAATTAAGTACTATCATTTTGGAGCA 120

QY 121 TAAAGGAGGAGAGATTTCAGCAAAACAGTGTCTTACAACTGAAAAACAAGTTAACTAA 180

DB 121 TAAAGGAGGAGAGATTTCAGCAAAACAGTGTCTTACAACTGAAAAACAAGTTAACTAA 180

QY 181 GTGACCCCCCTCTTGACAAAGATCAATGCCACAGTTAGCTTACGCCACATCATCA 240

DB 181 GTGACCCCCCTCTTGACAAAGATCAATGCCACAGTTAGCTTACGCCACATCATCA 240

QY 241 TGTAAATGCTTCTCGATTAAGCTTGTATTAATCTCTTTGCAAAAGCTTGCTACTTA 300

DB 241 TGTAAATGCTTCTCGATTAAGCTTGTATTAATCTCTTTGCAAAAGCTTGCTACTTA 300

QY 301 CCAGAAGTCTGCTACAGACAGAAAGATGGCATTTA 336

DB 301 CCAGAAGTCTGCTACAGACAGAAAGATGGCATTTA 336

RESULT 2
ADL90127
ID ADL90127 standard; DNA; 2381 BP.

AC ADL90127;

DT 20-MAY-2004 (first entry)

DE Chicken intestinal fatty acid binding protein, IFABP, gene, 5' region.

XX Chicken; de; intestinal fatty acid binding protein; IFABP;

KW gut specific promoter; transgenic.

OS Gallus gallus.

XX US2003177516-A1.

XX 18-SEP-2003.

XX 14-MAR-2002; 2002US-00099663.

XX 14-MAR-2002; 2002US-00099663.

XX (HORS/) HORSEMAN N D.

XX (PRAT/) PRATT S L.

PI Horseman ND, Pratt SL;

XX WPI; 2003-898653/82.

XX New nucleic acid molecule comprising an isolated avian gut-specific gene

PT expression control region, useful for regulating heterologous nucleic

XX acids in transgenic avians, and for generating transgenic birds.

PS Claim 1; SEQ ID NO 1; 28pp; English.

XX The invention relates to an isolated nucleic acid comprising an isolated

CC avian gut-specific gene expression control region appearing as

CC ADL90127 (Chicken intestinal fatty acid binding protein, IFABP, gene, 5'

CC region or ADL90128 (chicken IFABP promoter) or its degenerate variant.

CC Also included are a recombinant DNA molecule comprising an isolate avian

CC gut-specific gene expression control region operably linked to a nucleic

CC acid insert encoding a polypeptide, an expression vector that integrates

CC into a host cell (and comprising the isolated avian gut-specific gene

CC expression control region), expressing a heterologous polypeptide in a

CC host cell (by transfecting a eukaryotic cell with the recombinant DNA

CC molecule, and culturing the transfected cell in a medium suitable for

CC expression of a heterologous polypeptide under the control of an avian

CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression

CC control region encoded by the recombinant DNA molecule), a eukaryotic

CC cell transformed with the expression vector (or its progeny, which

CC expresses a heterologous polypeptide) and a transgenic avian having a

CC heterologous polynucleotide sequence comprising the nucleic acid insert.

CC The nucleic acids are useful for regulating heterologous nucleic acids in

CC transgenic avians, as probes in nucleic acid hybridisation assays for

CC detecting the IFABP gene expression control region, and for generating

CC transgenic birds. The present sequence is the Chicken intestinal fatty

CC acid binding protein, IFABP, gene, 5' region.

XX Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 336; DB 11; Length 2381;

XX Best Local Similarity 100.0%; Pred. No. 3.8e-87; Mismatches 0; Gaps 0;

XX Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTATTTATTTTCATTAGATAGCCGGTTTTTTTACTACAACTCAATTAAGATGAACAGATG 60

DB 1301 ATTATTTATTTTCATTAGATAGCCGGTTTTTTTACTACAACTCAATTAAGATGAACAGATG 1360

QY 61 AATGGTTTGTGACCTGTTTATAAGAGATTAATTAAGTACTATCATTTTGGAGCA 120

DB 1361 AATGGTTTGTGACCTGTTTATAAGAGATTAATTAAGTACTATCATTTTGGAGCA 1420

QY 121 TAAAGGAGGAGAGATTTCAGCAAAACAGTGTCTTACAACTGAAAAACAAGTTAACTAA 180

DB 1421 TAAAGGAGGAGAGATTTCAGCAAAACAGTGTCTTACAACTGAAAAACAAGTTAACTAA 1480

QY 181 GTGACCCCCCTCTTGACAAAGATCAATGCCACAGTTAGCTTACGCCACATCATCA 240

DB 1481 GTGACCCCCCTCTTGACAAAGATCAATGCCACAGTTAGCTTACGCCACATCATCA 1540

QY 241 TGTAAATGCTTCTCGATTAAGCTTGTATTAATCTCTTTGCAAAAGCTTGCTACTTA 300

DB 1541 TGTAAATGCTTCTCGATTAAGCTTGTATTAATCTCTTTGCAAAAGCTTGCTACTTA 1600

QY 301 CCAGAAGTCTGCTACAGACAGAAAGATGGCATTTA 336

DB 1601 CCAGAAGTCTGCTACAGACAGAAAGATGGCATTTA 1636

RESULT 3
ABL18616
ID ABL18616 standard; DNA; 2933 BP.

AC ABL18616;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 7321.

XX	Drosophila; developmental biology; cell signalling; insecticide;
KM	pharmaceutical; gene; db.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
XX	27-SEP-2001.
PF	23-MAR-2001; 2001WO-US009231.
XX	
PR	23-MAR-2000; 2000US--0191637P.
PR	11-UL-2000; 2000US-00614150.
PA	(PEXE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
XX	interactions.
PS	Claim 1; SEQ ID NO 7321; 21bp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC	sequences (ABLI01840-ABLI6175) and the encoded proteins (ABBS7737-
CC	ABBS72072). The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pcc_sequences
XX	
SO	Sequence 2933 BP; 883 A; 628 C; 606 G; 816 T; 0 U; 0 Other;
Query Match	11.6%; Score 39; DB 4; Length 2933;
Best Local Similarity	52.1%; Pred. NO. 0.75;
Matches 87; Conservative	0; Mismatches 80; Indels 0; Gaps 0;
OY	73 ACTGTTTAAAGAAGAGTAATTAAGATCATCATCTATTGAGGCAATAAGGAGGAG 132
DB	2766 ACATTAAATGACGATCTTACTTAACGAATTAATGATTATGATGCGAAGAGAGAT 2825
OY	133 AGATTACGAACAAGTGCTGTCTTACAAAGTGAACCAAGTAACTAAAGTACCCCTC 192
DB	2826 ATATTAAACCAAAATTTCTTATCGATGTGAACCAAGTGAACCGAATTTGATTCATC 2885
OY	193 CTTGACAAGATCAATGCCACAGTTGAGTTAGCCAGCAATCATC 239
DB	2886 CGTGCCCACTGAATGCAATCATCTTAGTCACACCAACCAATCAACCC 2932
RESULT 4	
ABL06564	
ID	ABL06564 standard; cDNA; 2927 BP.
XX	
AC	ABL06564;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 14174.
XX	
KM	Drosophila; developmental biology; cell signalling; insecticide;
KM	pharmaceutical; gene; ss.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.

XX	27-SEP-2001.
PD	
XX	
PF	23-MAR-2001; 2001WO-US009231.
XX	
XX	
PR	23-MAR-2000; 2000US-0191637P.
XX	
PR	11-JUL-2000; 2000US-00614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI, 2001-656860/75.
XX	
XX	P-PSDB; ABB62461.
PT	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signaling and cell-cell
PT	interactions.
PS	
PS	Claim 1; SEQ ID NO 14174; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC	sequences (AB101840-AB116175) and the encoded proteins (AB57737-
CC	AB572072). The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from Wipo at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 2927 BP; 882 A; 623 C; 606 G; 816 T; 0 U; 0 Other;
XX	
Query Match	11.1%; Score 37.2; DB 4; Length 2927;
Best Local Similarity	51.9%; Pred. No. 2.5;
Matches	84; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY	73 ACTGTTTATAAAGAAGTAAATAAGTACTATCATCTATTGAGCAATAGGAGGGAG 132
DB	2766 ACATTAAATGACGATCTTACTTAACGAATATAGATTATGATGCGAAGAGAGAT 2825
QY	133 AGATTTCAGCAAAACAGTGTGCTTACACAGTGGAAAAAGTTAACTAAATGACCCCTC 192
DB	2826 ATATTAAACGAATAATTCTTATCGATGTAAGAACAGTGAACGAATAATTGCATTCATC 2885
QY	193 CTGACAGATCATCATGCCACAGTTGAGCTTTAGCCGCCACA 234
DB	2886 CGTGGCCCACTGAATGCATCATCTTAATCGACACACCCATA 2927
XX	
RESULT 5	
ADJ79962/c	
ID	ADJ79962 standard; DNA; 123526 BP.
XX	
AC	ADJ79962;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Human glioma-associated oncogene-3 related DNA, SEQ ID No 11.
XX	
XX	glioma-associated oncogene-3; GAO3; cytosolic; developmental disorder;
KM	Grig1's cephalopolysyndactyly; Pallister-Hall syndrome;
KM	post-axial polydactyly; holoprosencephaly; Rubenstein-Teybi syndrome;
KM	basal cell nevroid syndrome; hyperproliferative disorder; cancer; human;
XX	ds.
XX	
OS	Homo sapiens.
XX	
PN	WO2003008549-A2.
XX	
PD	30-JAN-2003.
XX	

PF	15-JUL-2002; 2002W0-US022630.
XX	
PR	18-JUL-2001; 2001US-00910185.
PA	(ISIS-) ISIS PHARM INC.
XX	
P1	Bennett FC, Freiler SM;
XX	
DR	WPI; 2003-239322/23.
XX	
PT	New antisense oligonucleotides targeted to a nucleic acid encoding glioma-associated oncogene-3, useful for treating developmental disorders (e.g. holoprosencephaly) and hyperproliferative disorders (e.g. cancer).
XX	
PS	Disclosure; SEQ ID NO 11, 175pp; English.
CC	The invention relates to a novel compound 8-50 nucleobases in length
CC	targeted to a nucleic acid encoding glioma-associated oncogene-3 (GAO3)
CC	or a splice variant of GAO3. The novel compound specifically hybridizes
CC	with and inhibits the expression of GAO3 or its splice variant, or
CC	specifically hybridizes with an 8-nucleobase portion of an active site on
CC	a nucleic acid encoding GAO3. The antisense compound has cytosstatic
CC	activity. The antisense compound is useful for treating a disease or
CC	condition associated with glioma-associated oncogene-3 (GAO3), such as a
CC	developmental disorder including Greig's cephalopolysyndactyly, Pallister-
CC	Hall syndrome, post-axial polydactyly, holoprosencephaly, Rubenstein-
CC	Teysh syndrome or basal cell nevold syndrome, and a hyperproliferative
CC	disorder, such as cancer. This polynucleotide represents a DNA sequence
CC	relating to the human glioma-associated oncogene-3 (GAO3) of the
CC	invention.
SQ	Sequence 123526 BP; 34368 A; 24882 C; 26026 G; 38250 T; 0 U; 0 Other;
Query Match	11.1%; Score 37.2; DB 10; Length 123526;
Best Local Similarity	53.4%; Pred. No. 9;
Matches	78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
OY	12 CATTCAGATGCGCGGTCTTTTACTACACTCAATAAAGTAGACAAATGGCTTAAGT 71
DB	107264 CATTTCCCAATGCCCTTTAATTAAATGCTACGAAAATGCACGAAGAAGAGGGGTGCT 107205
OY	72 GACTGTTTTAAGAAGATTAAGATCATCTCATCTTGAGGCAATTAAGGGAGGGA 131
DB	107204 GTGTACCCTTTAAATTAATGACATGAAGAAAAAGAAAGCAGAGAGGAGGGA 107145
OY	132 GAGATTACGAAACACAGTGTGCTTACA 157
DB	107144 GGGAGAAAACCGAAAGCATCATGCA 107119
RESULT 6	
AAS58239	
ID	AAS58239 standard; cDNA; 403 BP.
XX	
AC	AAS58239;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	cDNA #915 encoding portion of a human colon tumour protein.
XX	
KM	Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.
OS	Homo sapiens.
XX	
PN	WO200173027-A2.
XX	
PD	04-OCT-2001.
XX	
PF	22-MAR-2001; 2001WO-US009246.
XX	
PR	24-MAR-2000; 2000US-0191597P.
PR	04-MAY-2000; 2000US-0202024P.
PR	05-MAY-2000; 2000US-0202189P.

XX	(CORI-) CORIXA CORP.
XX	
F1	Meagher MJ, Xu J, King GE;
XX	
DR	WPI; 2001-611627/70.
XX	
PT	New colon tumor proteins and related nucleic acid, useful for treatment,
XX	
PS	prevention, diagnosis and monitoring of cancer.
XX	
PS	Claim 4; Page 202; 299pp; English.
XX	
CC	Th present invention relates to the isolation of novel cDNA sequences
CC	encoding for at least an immunogenic portion of human colon tumour
CC	proteins. The sequences of the invention are useful in pharmaceutical
CC	compositions and vaccines for the prevention and treatment of cancers
CC	such as colon cancer. They are also useful for the diagnosis and
CC	monitoring of such cancers. Antibodies to the colon tumour proteins and
CC	antigen presenting cells that express polynucleotides encoding colon
CC	tumour proteins can be used to inhibit the development of cancers. T-
CC	cells that react specifically with colon tumour proteins are useful for
CC	removing tumour cells from samples (e.g. blood) and for cancer treatment.
CC	The polynucleotides sequences are also useful in gene therapy. AAS57325-
CC	AAS5880 represent the cDNA sequences of the invention that encode for
CC	portions of human colon tumour proteins
SQ	
Sequence	403 BP; 136 A; 63 C; 88 G; 105 T; 0 U; 11 Other;
Query Match	10.9%; Score 36.6; DB 4; Length 403;
Best Local Similarity	50.6%; Pred. No.1.9; Mismatches 82; Indels 0; Gaps 0
Matches 84; Conservative	0; Mismatches
Qy	26 TTTTCTACTCACTCAATTAAGATGACAGAAATGATGGTTAGTGACTGTTTATAAG 85
Db	178 TGTTTAAAGTTAAAGAAAAAGAGCTGCAGAGTATTATATAAACTGTCCTTTAGAAAAA 237
Qy	86 AAGAGTATAAAGATCTATCATCTTTGAGGCAATAAGGGGAGAGATTCAGCAAC 145
Db	238 AACAGCAGGAAGACCATTTTACCATATGAATGGAAAAAGGAAGAACTATTATAGAAAC 297
Qy	146 AGTGTGCTTACAGATGAAAAACAAGTTTAACTTAAAGTACCCTCT 191
Db	298 TTTCGCTGATTAAAAAATTTTAAAAAATTTTGGCAGAAACCCCT 343
RESULT 7	
ABLI0846	
ID	ABLI0846 standard; cDNA; 3322 BP.
XX	
AC	ABLI0846;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 27020.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
XX	
KW	pharmaceutical; gene; ss.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US009231.
XX	
PR	23-MAR-2000; 2000US-0191637P.
XX	
PR	11-JUL-2000; 2000US-00614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	

DR WPI; 2001-656860/75.
DR P-PSDB; ABB66743.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 27020; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3322 BP; 993 A; 715 C; 742 G; 872 T; 0 U; 0 Other;
XX
Query Match 10.9%; Score 36.6; DB 4; Length 3322;
Best Local Similarity 57.4%; Pred. No. 3.9;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
XX
QY 1 ATTTATTATTTTCATTAGATAGCCGCTTTTACTACAACTCAATAAGATGAAGATG 60
DB 3191 ATTGAGAAATTTCTTTTATCTGTTTGTATATCTACACAAATTTGGTCCAGTAGA 3250
XX
QY 61 AATGGGTAGTACTGCTGTTTAAAGAGTAATTAAGATCAATCATCTATTGGA 115
DB 3251 AGTTAGATTTTAATTCGGTTAGCAGAGTCAATTAACAATTAATTAATCACTTGA 3305
XX
RESULT 8
AAK20251/C
ID AAK20251 standard; DNA; 53585 BP.
XX
AC AAK20251;
XX
DT 04-MAY-1999 (first entry)
XX
DE *Borrelia burgdorferi* polynucleotide sequence #4.
XX
KW *Borrelia burgdorferi*; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
XX
OS *Borrelia burgdorferi*.
XX
PN WO9858943-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US012764.
XX
XX 20-JUN-1997; 97US-0050359P.
XX 22-JUL-1997; 97US-0053344P.
XX 22-JUL-1997; 97US-0053377P.
XX 03-SEP-1997; 97US-0057483P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMUNE INC.
XX
PI Frazer C, White OR, Clayton R, Dougherty BA, Lathigra R;
PI Smith HO;
XX
DR WPI; 1999-081217/07.
XX
PT New isolated *Borrelia burgdorferi* nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention and
PT therapy of infections, particularly Lyme disease.

XX
PS Claim 1; Page 801-831; 1128pp; English.
XX
CC AAK20248 to AAK20402 represent polynucleotide sequences isolated from
CC *Borrelia burgdorferi* (Bb). Products derived from Bb can be used for the
CC detection, diagnosis, characterisation, prevention and therapy of Bb
CC infections, e.g. Lyme disease. They can also be used for the production
CC of biosynthetic products, e.g. enzymes. *Borrelia* belongs to a family of
CC motile, spiral-shaped bacteria called Spirochetes. Spirochetes are
CC pathogenic in humans and *Borrelia* causes epidemic and endemic relapsing
CC fever, and Lyme borreliosis, more commonly known as Lyme disease
XX
SQ Sequence 53585 BP; 19235 A; 8081 C; 7001 G; 19266 T; 0 U; 2 Other;
XX
Query Match 10.9%; Score 36.6; DB 2; Length 53585;
Best Local Similarity 58.9%; Pred. No. 10;
Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
XX
QY 78 TTATTAAGAAGAGTAATTAAGATCAATCTTTGAGGCAATTAAGGAGGAGAGATT 137
DB 44978 TTTAAAAATGTTATTAATTAAGATTACTCACTTATATCAAGCATTAATGATGAGAGATG 44919
XX
QY 138 CAGCAACAGTGTGCTTACAGTGAAACCAAGTTAACTAAGTGA 184
DB 44918 CAGCATCAATTGAACATTTTAAGAAAGAGAAATGAATCAATATGTCA 44872
XX
RESULT 9
AAH53268
ID AAH53268 standard; DNA; 1587 BP.
XX
AC AAH53268;
XX
DT 03-SEP-2001 (first entry)
XX
DE *S. epidermidis* open reading frame nucleotide sequence SEQ ID NO:1929.
XX
KW *Staphylococcus epidermidis* SRI strain; infection; diagnosis; vaccination;
KW endocarditis; ds.
XX
OS *Staphylococcus epidermidis*.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US030782.
XX
PR 09-NOV-1999; 99US-0164258P.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Kimerly WJ;
XX
XX WPI; 2001-316495/33.
XX P-PSDB; AAG82418.
XX
PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
PT useful for vaccinating against infections, e.g. endocarditis.
XX
PS Claim 8; Page 527-528; 2188pp; English.
XX
AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the *S.*
CC *epidermidis* polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to

AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX Sequence 1587 BP; 617 A; 268 C; 201 G; 501 T; 0 U; 0 Other;

Query Match 10.8%; Score 36.2; DB 4; Length 1587;
Best Local Similarity 53.1%; Pred. No. 3.9;

Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 141 CAACAGTGTGCTTACAGTGAACAAAGTTAACTAAAGAGACCCCTCTTGACA 200

DB 203 CAGACGATGACTTTAAAGTTGAACCTTAATTTTACAGAGTTACCTCAAGCCCATACTA 262

QY 201 GATCAATGCGACAGTTAGCTTACGACCATCATCATGTAATTCCTTCCGATA 260

DB 263 GAAATATTAACAAATTAATTCACAAAGCATATGATATCATCAATTACACTTGTGAAA 322

QY 261 AGCCTGTCATTAATTCCTTTGCA 285

DB 323 ATAAATACACAAATTACTTGTGCA 347

RESULT 10

ABN92204 ID ABN92204 standard; DNA; 1878 BP.

XX ABN92204;

DT 24-JUL-2002 (first entry)

XX Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:1667.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

XX antibacterial; gene therapy; gene; ds.

XX Staphylococcus epidermidis.

XX US6380370-B1.

XX 30-APR-2002.

XX 13-AUG-1998; 98US-00134001.

XX 14-AUG-1997; 97US-0055779P.

XX 08-NOV-1997; 97US-0064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2002-381255/41.

XX P-PSDB; ABP39659.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermis

XX PT polypeptide, useful for diagnosing and treating bacterial infections.

XX Disclosure; SEQ ID NO 1667; 267bp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading

XX CC frame (ORF) nucleic acid sequences which encode the amino acid sequences

XX CC given in ABP35124 to ABP37960. The S. epidermidis sequences have

XX CC antibacterial activity and can be used in gene therapy. The sequences can

XX CC also be used in the diagnosis and treatment of bacterial infections,

XX CC particularly S. epidermidis infections. The sequences can be used to

XX CC screen for compounds able to interfere with the S. epidermidis life cycle

XX CC or inhibit S. epidermidis infection. N.B. The sequence data for this

CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site

XX Sequence 1878 BP; 734 A; 304 C; 239 G; 601 T; 0 U; 0 Other;

Query Match 10.8%; Score 36.2; DB 6; Length 1878;
Best Local Similarity 53.1%; Pred. No. 4.2;

Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 141 CAACAGTGTGCTTACAGTGAACAAAGTTAACTAAAGAGACCCCTCTTGACA 200

DB 494 CAGACGATGACTTTAAAGTTGAACCTTAATTTTCAAGAGTTACCTCAAGCCCATACTA 553

QY 201 GATCAATGCGACAGTTAGCTTACGACCATCATCATGTAATTCCTTCCGATA 260

DB 554 GAAATATTAACAAATTAATTCACAAAGCATATGATATCATCAATTACACTTGTGAAA 613

QY 261 AGCCTGTCATTAATTCCTTTGCA 285

DB 614 ATAAATACACAAATTACTTGTGCA 638

RESULT 11

ADSO1975 ID ADSO1975 standard; DNA; 1878 BP.

XX ADSO1975;

DT 04-NOV-2004 (first entry)

XX Staphylococcus epidermis polynucleotide seqid 1270.

XX antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;

XX recombinant expression vector; infection; computer readable medium;

XX computer based system; gene; ds.

XX Staphylococcus epidermidis.

XX US2004147734-A1.

XX 29-JUL-2004.

XX 01-DEC-2003; 2003US-00724972.

XX 08-NOV-1997; 97US-0064964P.

XX 13-AUG-1998; 98US-00134001.

XX 29-NOV-1999; 99US-00450969.

XX (DOUC/) DOUCETTE-STAMM L.

XX (BUSH/) BUSH D.

XX Doucette-Stamm L, Bush D;

XX WPI; 2004-580138/56.

XX P-PSDB; ADS05747.

XX New isolated polypeptide and encoding nucleic acid derived from

XX PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or

XX PT treating an S. epidermidis bacterial infection.

XX Claim 5; SEQ ID NO 1270; 741bp; English.

XX The invention describes an isolated nucleic acid comprising a nucleotide

XX CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:

XX CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any

XX CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as

XX CC given in the specification. Also described are: a recombinant expression

XX CC vector; a cell comprising a recombinant expression vector of (1);

XX CC producing an S. epidermidis polypeptide; an isolated nucleic acid

XX CC comprising a nucleotide sequence of at least 8 nucleotides in length; a

XX CC vaccine composition for prevention or treatment of an S. epidermidis

XX CC infection, comprising a nucleic acid cited above and a carrier; treating

XX CC a subject for S. epidermidis infection; a recombinant or substantially

CC epidermidis polypeptides (II) via the production of vectors containing

CC and (II) can have antibacterial activity and therefore can be used in

CC vaccination. The nucleic acids (I) may be used to produce the S.

CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464

XX SQ Sequence 3368 BP; 1163 A; 447 C; 569 G; 1189 T; 0 U; 0 Other;

XX Query Match 10.8%; Score 36.2; DB 4; Length 3368;
XX Best Local Similarity 53.1%; Pred. No. 5.1;
XX Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 141 CAACAGCTGTCTTACAGAGTGAACCAAGTTAAAGTACCCCTCTTGACAA 200
DB 1038 CAGACCATGACTTTAAAGTTGAACCTTATTACAGAGTTACCTCAAGCCCATCTA 979

QY 201 GATCATGCGACAGCTTGAGCTTAGCCAGCCATCATCATGTAATGCTTCCGATA 260
DB 978 GAAATATTAACCAATATCTCAAGCATATGATATCATCATTAATCACTTGAAA 919

QY 261 AGCCTGTCATAATCTCTTTGCA 285
DB 918 ATTAATATCAACAATTAATTCTTGCA 894

XX RESULT 14
XX ID AAX84331 standard; DNA; 6088 BP.
XX AAX84331;
XX 08-SEP-1999 (first entry)
XX Stealth virus nucleic acid clone, SEQ ID NO: 23.
XX Stealth virus; detection; diagnosis; infection; ss.
XX Stealth virus.
XX OS
XX Key Location/Qualifiers
XX FT misc_difference 5841
XX FT /*tag= a
XX FT /note= "this nucleotide is represented as a * in the
XX FT specification, and is included to maintain the base
XX FT numbering given in the specification"
XX FT 5997. .8133
XX FT /*tag= b
XX FT /note= "this nucleotide is represented as a * in the
XX FT specification, and is included to maintain the base
XX FT numbering given in the specification"
XX FT misc_difference 8275
XX FT /*tag= d
XX FT /note= "this nucleotide is represented as a * in the
XX FT specification, and is included to maintain the base
XX FT numbering given in the specification"
XX WO9934019-A1.
XX PN
XX 08-JUL-1999.
XX PD
XX 30-DEC-1998; 98WO-US027744.
XX PF
XX

PR 30-DEC-1997; 97US-00001184.
XX
XX PA (MART/) MARTIN W J.
XX
XX PI Martin WJ;
XX
XX DR WPI; 1999-405521/34.
XX
XX PT Novel strains of stealth virus.
XX
XX PS Claim 19; Page 64-66; 95pp; English.
XX
XX CC This sequence represents a Stealth virus nucleic acid clone. The
XX CC invention relates to a method of detecting and characterising a stealth
XX CC virus by reacting a sample suspected of containing a stealth virus with a
XX CC probe from a known stealth virus and sequencing the resultant isolated
XX CC nucleotide. The method comprises the steps of: (a) isolating DNA or RNA
XX CC from a sample suspected of containing a stealth virus, e.g. a culture of
XX CC cells showing a viral cytopathic effect; (b) testing the reactivity of
XX CC the isolated DNA or RNA with a molecular probe that contains at least 18
XX CC or more contiguous nucleotides identical to sequence previously
XX CC identified from a stealth virus; and, optionally (c) sequencing the
XX CC isolated DNA or RNA molecules that react with the probe. The method is
XX CC used to detect stealth virus in a biological product, food or in the
XX CC environment. The method is also used to evaluate agents for their
XX CC inhibitory or stimulatory effects on stealth virus replication and to
XX CC determine capacity of the virus to recombine with and potentially alter
XX CC the nucleic acid sequences of a cell or bacterium

XX SQ Sequence 6088 BP; 1577 A; 1455 C; 1442 G; 1427 T; 0 U; 187 Other;

XX Query Match 10.8%; Score 36.2; DB 2; Length 6088;
XX Best Local Similarity 20.1%; Pred. No. 6.3;
XX Matches 43; Conservative 77; Mismatches 94; Indels 0; Gaps 0;

QY 42 AATAAGATGACAGATGATGAGTGTAGTACTGTTATTAAGAAGTAAAGATA 101
DB 5824 AATTGAAGAATCTTCHANYWVTSRMMNMAMHADAATAGRAHNRHMNATNATNST 5883

QY 102 CTATCATCTTTGAGCATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTACAGTG 161
DB 5884 ATHBDCMBGTMSNRMANYSVBARAHBCMHNMARTNMHNRTHDITBCEHNMARTN 5943

QY 162 GAAACCAAGTTAACTAAAGTACCCCTCTTACAGATCAATGACAGCTGAGCT 221
DB 5944 NEMAMHNMARTNMGSTWRDMMDHCCDBCTTGDANBCDAADRTNRYTABNRDDCM 6003

QY 222 TTAGCCAGCCATCATCATGTAATGCTTTC 255
DB 6004 NTRSMWARYNRWATNDCMWTSMWARYNRWATB 6037

XX RESULT 15
XX ID ABRK1354 standard; DNA; 8648 BP.
XX ABRK1354;
XX 23-APR-2002 (first entry)
XX Signal transduction associated gene modified DNA #99.
XX DE
XX Human; signal transduction associated gene; cytosine methylation state;
XX Cpg island; signal transduction associated disease; solid tumour; cancer;
XX antitumour; cytosolic; mutant; de.
XX OS Homo sapiens.
XX OS Synthetic.
XX WO200200926-A2.
XX PN
XX 03-JAN-2002.
XX PD
XX

PF 29-JUN-2001; 2001WO-EP007472.

XX

PR 30-JUN-2000; 2000DE-01032529.

PR

01-SEP-2000; 2000DE-01043826.

XX

XX

PA

(EPIG-) EPIGENOMICS AG.

PI

Olek A, Piepenbrock C, Berlin K;

XX

DR

WPI; 2002-147896/19.

XX

PT

Oligonucleotide for diagnosis and therapy of diseases associated with

signal transduction e.g. cancer, comprises chemically modified genomic

sequences of genes associated with signal transduction.

XX

PS

Claim 1; SEQ ID NO 197; 24pp; English.

XX

CC

The present invention relates to chemically modified DNA sequences of

signal transduction associated genes. The DNA sequences are chemically

modified using a solution of bisulphite, hydrogen sulphite or disulphite.

CC

Also disclosed are oligonucleotides and/or PNA oligomers for detecting

the cytosine methylation state (Cpg islands) of these genes, and a method

for the diagnosis and/or therapy of genetic and epigenetic parameters of

genes associated with signal transduction. The genomic DNA can be

obtained from cells or cellular components which contain DNA, e.g. cell

lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,

tissue embedded in paraffin such as tissue from eyes, intestine, kidney,

brain, heart, prostate, lung, breast or liver, histologic object slides,

and all their possible combinations. The sequences of the invention are

useful for the diagnosis and therapy of diseases associated with signal

transduction e.g. solid tumours and cancer. ABK3158-ABK31545 represent

chemically pretreated genomic DNA sequences of different genes associated

with signal transduction, or their complementary sequences. Note: The

sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from the

European Patent Office

CC

XX

Seq

Sequence 8648 BP; 2149 A; 236 C; 1917 G; 4346 T; 0 U; 0 Other;

Query Match

Best Local Similarity 10.8%; Score 36.2; DB 6; Length 8648;

Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY

3 TATTTATTTTCATTAGATGCGGTTTCTACTACAACTCAATAGATGACAGAAATGAA 62

DB

940 TTTTGGTTTATACGTCGTGTTAGTTAGTATTATGATTAAGTTAGTTTGATAT 999

QY

63 TGGGTTAGACCTGTTTATTAAGAAAGATTAATTAAGTACTATCATCATTTGA 115

DB

1000 TTGGGAAGTGAATGTTTATGAGATTAATTAATTAATTTGTTATTTGA 1052

Search completed: November 21, 2005, 05:53:48
Job time : 271.264 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 02:06:40 ; Search time 1583.72 Seconds
(without alignments)
10280.216 Million cell updates/sec

Title: US-10-099-663-2

Perfect score: 336
Sequence: 1 attattatttcattcatgata.....agacagaaagatgcattta 336

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*\n2: gb_hcg:*\n3: gb_in:*\n4: gb_cm:*\n5: gb_ov:*\n6: gb_pat:*\n7: gb_ph:*\n8: gb_pl:*\n9: gb_pr:*\n10: gb_ro:*\n11: gb_scs:*\n12: gb_sy:*\n13: gb_un:*\n14: gb_v1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	42.8	12.7	88688	9	AC087429	AC087429 Homo sapi
C 2	42.8	12.7	187003	2	AC023353	AC023353 Homo sapi
C 3	41.8	12.4	252420	3	AE014841	AE014841 Plasmodiu
C 4	41.4	12.3	1663	3	AK116155	AK116155 Clona int
C 5	41.4	12.3	99855	8	OSJN00175	OSJN00175 Oryza sat
C 6	41.4	12.3	121130	8	OSJN00292	OSJN00292 Oryza sat
C 7	40.6	12.1	209024	10	AL928605	AL928605 Oryza sat
C 8	40.4	12.0	60756	2	AL583823_3	AL583823_3 Mouse DNA
C 9	40.4	12.0	131708	2	AL354710	AL354710 Human DNA
C 10	40.4	12.0	174294	2	AC026155	AC026155 Homo sapi
C 11	40.2	12.0	110000	10	AE014175_2	AE014175_2 Contamination (3 of
C 12	40.2	12.0	201783	10	AC114410	AC114410 Mus muscu
C 13	39.8	11.8	177655	2	CR318618	CR318618 Danio rer
C 14	39.8	11.8	217375	5	BX537337	BX537337 Zebrafish
C 15	39.6	11.8	162605	5	AC068698	AC068698 Homo sapi
C 16	39.6	11.8	164275	9	AC005157	AC005157 Homo sapi
C 17	39.6	11.8	230585	10	AC110259	AC110259 Mus muscu
C 18	39.4	11.7	393	8	YSCMTORIE	K02992 Yeast (S.ce
C 19	39.4	11.7	396	8	MIECORIS	X01216 Yeast mtloc

C 20	39.4	11.7	130574	9	AL160174	AL160174 Human DNA
C 21	39.4	11.7	141668	10	BX001028	BX001028 Mouse DNA
C 22	39.4	11.7	172945	9	AC067719	AC067719 Homo sapi
C 23	39.4	11.7	193306	2	AC025512	AC025512 Homo sapi
C 24	39.4	11.7	194638	2	AC112913	AC112913 Homo sapi
C 25	39.4	11.7	238301	2	AC094376	AC094376 Rattus no
C 26	39	11.6	2933	6	CQ597407	CQ597407 Sequence
C 27	39	11.6	9124	2	AC020330	AC020330 Drosophi1
C 28	39	11.6	86258	5	BX842699	BX842699 Zebrafish
C 29	39	11.6	108924	3	AC005889	AC005889 Drosophi1
C 30	39	11.6	177997	3	AC092492	AC092492 Drosophi1
C 31	39	11.6	270775	3	AE003624	AE003624 Drosophi1
C 32	38.8	11.5	1031	3	TEH152A1A	TEH152A1A Tetrahymena
C 33	38.8	11.5	109465	8	AP006106	AP006106 Lotus cor
C 34	38.8	11.5	140266	10	AC124406	AC124406 Mus muscu
C 35	38.8	11.5	210359	9	HS4430K20	HS4430K20 Human DNA
C 36	38.8	11.5	212445	10	AC108419	AC108419 Mus muscu
C 37	38.6	11.5	165873	5	AL954323	AL954323 Zebrafish
C 38	38.6	11.5	207372	10	AL731851	AL731851 Mouse DNA
C 39	38.6	11.5	214950	2	AC110406	AC110406 Rattus no
C 40	38.6	11.5	246528	2	AC128573	AC128573 Rattus no
C 41	38.4	11.4	66569	2	AC103837	AC103837 Homo sapi
C 42	38.4	11.4	154433	10	AL732447	AL732447 Mouse DNA
C 43	38.4	11.4	166860	10	AC115924	AC115924 Mus muscu
C 44	38.4	11.4	177967	10	AC122816	AC122816 Mus muscu
C 45	38.4	11.4	189019	2	AC129426	AC129426 Rattus no

ALIGNMENTS

RESULT 1	AC087429	88688 bp	DNA	linear	PRI 11-APR-2002
LOCUS	AC087429/c				
DEFINITION	Homo sapiens chromosome 3 clone 2063K18 map 3p, complete sequence.				
ACCESSION	AC087428				
VERSION	AC087429.2	GI:20136692			
KEYWORDS	HTG.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
REFERENCE	Bukacinski, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 88688)				
AUTHORS	Xiong, H., Zhou, Y., Dong, H., Lian, W., Chen, B., Zhang, C., Zhang, Y., Cai, Z., Ying, H., Wang, H., Gu, W., Zhu, G., Tu, Y., Zhang, X., Jia, J., Shen, H., Zhang, D., Wu, C., Lu, G., Zhong, M., Jiang, H., Ren, S., Fu, G., Chen, Z. and Huang, M.				
TITLE	Chromosome 3p genomic sequence				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 88688)				
AUTHORS	Liu, B., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Guan, J., Guan, Q., Gu, X., Guo, D., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Luo, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, W., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-JAN-2001) 1. Chinese Human Genome Center at Shanghai				
REFERENCE	2. Institute of Genetics, Chinese Academy of Sciences, P.R.China				
AUTHORS	3 (bases 1 to 88688)				
	Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Guan, J., Guan, Q., Gu, X., Guo, D., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Luo, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, W., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.				

TITLE
JOURNAL
Zhang, X., Zhang, X., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.
Direct Submission
Submitted (11-APR-2002) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing,
100101, P.R.China
On Apr 11, 2002 this sequence version replaced gi:12039240.

COMMENT

-----Genome Center
Center: Beijing Center
Website: <http://hgclib.ac.cn>
<http://www.genomics.org.cn>
Contact: hgclib.ac.cn
-----Project Information
Center project name: 1% project
Center clone name: 2063K18
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: ET 55% of reads
Chemistry: Dye-terminator: Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 89950 bases at least Q40
Consensus quality: 90370 bases at least Q30
Consensus quality: 90443 bases at least Q20
Insert size: 88688; sum-of-contigs
Quality coverage: 7.47x in Q20 bases; sum-of-contigs

FEATURES

Location/Qualifiers
1..88688
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="2063K18"

ORIGIN

Query Match 12.7%; Score 42.8; DB 9; Length 88688;
Best Local Similarity 53.6%; Pred. No. 1.3;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1 ATTATATTTTCTAGATAGCCGGTTTCTTACTCACTCAATTAAGTGAACAGATG 60
DB 1637 ATTAACTTAACAAATTAAGCTTTATATCTCTTAACAACGAACATTAGAGAACT 1578
QY 61 AATGGCTGTGACCTGTTTATAAGAGCTATAAGATCATATCATTTGAGGCA 120
DB 1577 TAAAGATTAAGTCCATCTATTAATAGATTAAGAACTATATTAAGATTAAGCA 1518
QY 121 TAAAGAGGAGAGATTTCAGCAACAGTGTGCTTACAGTGAATAA 166
DB 1517 TAAAGACACTTAATCTAGAAATAATTAATAAAATGTACAA 1472

RESULT 2
AC023353/c 187003 bp DNA linear HTG 17-AUG-2000
LOCUS Homo sapiens chromosome 20 clone RP11-775C23, WORKING DRAFT
DEFINITION
SEQUENCE, 28 unordered pieces.
AC023353
AC023353 GI:9211518
VERSION
AC023353.3 GI:9211518
KEYWORDS
HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 187003)
REFERENCE
JOURNAL
The sequence of Homo sapiens clone
2 (bases 1 to 187003)
AUTHORS
Waterston, R.H.
TITLE
Direct Submission

JOURNAL

COMMENT

Submitted (11-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 15, 2000 this sequence version replaced gi:7235346.

-----Genome Center -----
Center: Washington University Genome Sequencing Center
Center: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
-----Project Information -----
Center project name: H_NH0775C23
-----Summary Statistics -----
Sequencing vector: M13; 57%
Sequencing vector: plasmid; 43%
Chemistry: Dye-terminator: Big Dye; 43% of reads
Chemistry: Dye-terminator: Big Dye; 43% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 172430 bases at least Q40
Consensus quality: 176603 bases at least Q30
Consensus quality: 178873 bases at least Q20
Insert size: 18900; agarose-fp
Insert size: 184303; sum-of-contigs
Quality coverage: 3.50 in Q20 bases; agarose-fp
Quality coverage: 3.66 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1408: contig of 1408 bp in length
1409
1508: gap of unknown length
1509
3342: contig of 1834 bp in length
3343
3442: gap of unknown length
3443
5329: contig of 1887 bp in length
5429: gap of unknown length
5430
8611: contig of 3182 bp in length
8612
8711: gap of unknown length
8712
11044: contig of 2333 bp in length
11045
11144: gap of unknown length
11145
15105: contig of 3961 bp in length
15106
15205: gap of unknown length
15206
17732: contig of 2527 bp in length
17733
17832: gap of unknown length
17833
21791: contig of 3959 bp in length
21792
21891: gap of unknown length
21892
23547: contig of 1656 bp in length
23548
23647: gap of unknown length
23648
28121: contig of 4474 bp in length
28122
28221: gap of unknown length
28222
34018: contig of 5797 bp in length
34019
34118: gap of unknown length
34119
38469: contig of 4351 bp in length
38470
38569: gap of unknown length
38570
44271: contig of 5702 bp in length
44272
44371: gap of unknown length
44372
51457: contig of 7086 bp in length
51458
51557: gap of unknown length
51558
58979: contig of 7422 bp in length
58979
59079: gap of unknown length
59080
66559: contig of 6580 bp in length
66560
65759: gap of unknown length
65760
74077: contig of 8318 bp in length
74078
74177: gap of unknown length
74178
82579: contig of 8402 bp in length
82579
82579: gap of unknown length
82580
92223: contig of 9544 bp in length
92223
92323: gap of unknown length
92324
96555: contig of 7332 bp in length
96555
99755: gap of unknown length

	*	99756	108570:	contig of 8815 bp in length
	*	108571	108670:	gap of unknown length
	*	108671	115388:	contig of 6718 bp in length
	*	115389	115488:	gap of unknown length
	*	115489	123778:	contig of 8290 bp in length
	*	123779	123878:	gap of unknown length
	*	123879	123927:	contig of 8449 bp in length
	*	123928	132427:	gap of unknown length
	*	132428	145195:	contig of 12768 bp in length
	*	145196	145295:	gap of unknown length
	*	145296	158332:	contig of 13037 bp in length
	*	158333	158432:	gap of unknown length
	*	158433	171052:	contig of 12620 bp in length
	*	171053	171152:	gap of unknown length
	*	171153	187003:	contig of 15851 bp in length
FEATURES			Location/Qualifiers	
	source		1..187003	
			/organism="Homo sapiens"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:9606"	
			/chromosome="20"	
			/clone="RP11-775C23"	
			1..1408	
misc_feature		/note="assembly_name:Contig29"	1509..3342	
misc_feature		/note="assembly_name:Contig30"	3443..5329	
misc_feature		/note="assembly_name:Contig31"	5430..8611	
misc_feature		/note="assembly_name:Contig32"	8712..11044	
misc_feature		/note="assembly_name:Contig33"	11145..15105	
misc_feature		/note="assembly_name:Contig34"	15206..17732	
misc_feature		/note="assembly_name:Contig35"	17833..21791	
misc_feature		/note="assembly_name:Contig36"	21892..23547	
misc_feature		/note="assembly_name:Contig37"	23648..28121	
misc_feature		/note="assembly_name:Contig38"	28222..34018	
misc_feature		/note="assembly_name:Contig39"	34119..38469	
misc_feature		/note="assembly_name:Contig40"	38570..44271	
misc_feature		/note="assembly_name:Contig41"	44372..51457	
misc_feature		/note="assembly_name:Contig42"	51558..58979	
misc_feature		/note="assembly_name:Contig43"	59080..65659	
misc_feature		/note="assembly_name:Contig44"	65760..74077	
misc_feature		/note="assembly_name:Contig45"	7418..82579	
misc_feature		/note="assembly_name:Contig46"	8260..92223	
misc_feature		/note="assembly_name:Contig47"	92324..99655	
misc_feature		/note="assembly_name:Contig48"	99756..108570	
misc_feature		/note="assembly_name:Contig49"	108671..115398	
misc_feature		/note="assembly_name:Contig50"	115489..123778	
misc_feature		/note="assembly_name:Contig51"	123879..132327	
misc_feature		/note="assembly_name:Contig52"	132428..145195	
misc_feature		/note="assembly_name:Contig53"	145296..158332	

```
misc_feature /note="assembly_name:Contig54"  
158433..171052  
/note="assembly_name:Contig55  
clone_end:77  
vector_end:right"  
171153..187003  
/note="assembly_name:Contig56"  
  
ORIGIN  
  
Query Match 12.7%; Score 42.8; DB 2; Length 187003;  
Best Local Similarity 53.6%; Pred.No.1.2;  
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
  
OY 1 ATATTATTTTCATTAGATGATGACCGGTCTTTTACTCAAACTCAATAATGATGAACAATG 60  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 119465 ATTACATTAACAAATTTAGTAGCCCTTTTATATCTCTTTTAAACAACGAACTTAGAACA 119406  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
OY 61 AATGGGTAGTGACGCCTTTATPAAAAGAAGTAATAAAGTACTATCATCATTGGAGCAA 120  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 119405 TTAAGGTATGTCACATGTAATATPAAATPAAAGATPAAAGAACCTATATATPAAAGATPAA 119346  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
OY 121 TAAGGAGGAGAGAGATTCAGCAAAACAGTGTGCTTCAAGTGGAAAA 166  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 119345 TTAAGACACTAATATCTTAGAAATTAATTTATATAAAATGTACAAA 119300  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
  
RESULT 3  
AE014841  
LOCUS Plasmodium falciparum 3D7 chromosome 11 section 6 of 8 of the  
DEFINITION complete sequence.  
ACCESSION AE014841 AE014186  
VERSION AE014841.1 GI:23496259  
KEYWORDS .  
SOURCE Plasmodium falciparum 3D7  
ORGANISM Plasmodium falciparum 3D7  
REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
1 (bases 1 to 252420)  
AUTHORS Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,  
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,  
James,K., Eiden,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,  
Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Sun,B., Peterson,J.,  
Angiloul,S., Perlea,M., Allen,J., Selengut,J., Haft,D.,  
Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H.,  
Fruminholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,  
Cummins,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,  
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and  
Barrell,B.  
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum  
JOURNAL Nature 419 (6906), 498-511 (2002)  
PUBMED 12368864  
REFERENCE 2 (bases 1 to 252420)  
AUTHORS Gardner,M.J.  
TITLE Direct Submission  
JOURNAL Submitted (13-SEP-2002) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
FEATURES  
source location/Qualifiers  
1..252420  
/organism="Plasmodium falciparum 3D7"  
/mol_type="genomic DNA"  
/iso_late="3D7"  
/db_xref="taxon:36329"  
/chromosome="11"  
complement(85..120)  
/note="AT-rich."  
repeat_region /rpt_type=tandem  
repeat_region /rpt_type=tandem  
repeat_region complement(170..253)  
/rpt_type=cadem  
/rpt_unit="(ta)n"  
224..253  
/rpt_type=tandem  
repeat_region /rpt_type=cadem  
360..399  
/rpt_type=tandem
```

	repeat__region	complement(360..399) /rpt_type=tandem /rpt_unit="(ta)n"
gene	<419..>.3287 /locus_tag="PF11_0334" join(<419..3121>,3279..>.3287)	
mRNA	/locus_tag="PF11_0334" join(419..3121,3279..3287)	
CDS	/locus_tag="PF11_0334" /codon_start=1 /product="hypothetical protein" /protein_id="AAW35918.1" /db_xref="GI:23496260" 	
	ttranslatio n="MGTPDNTDETYNITLHLEERBRKRNKL DVSHNKDPEFL KLYLVY NACLSANSTDI YQVEYNIRLSINSVNNCDYDOI KSCVI SNDDPL KYNYNFYNDP FVL EMDNETPTSSSYHLNEEMKEELSEFPIHHMKKI KEFEVGNDISRKRSPICCTEN VAMSDKTETENNNSLT SINNNNNNLSTTFOONSHLPFTLL EKRRKKKNKP LYLISHIGIKNESI TVNHNEYHTSNYL PYVP MKMNIF ESTCQSQHHSIS THMDD PFLCKOS INRMKLRSGNSGYFCMLQNIEAONDPYEMI TKYLDVT KKRSNIATD NVSLCSMDINTAKRKRRKKRI KDQRSNKREMYDKIN IYNNPINLTSNHSYSYEILSI NSINI IDEBNTLENYYIDMEI LYNTSNDNRKKDOITKLINEPTLI VOINKKKNI KR LSASVDYII KTMMINEKITKSNOQR IDRDKNMAL YOFPMSGASLSI YEYDL PLR RKFPLYILTGVIPCSI FYFI GLPIL OELSIOGL SGCILNIS FLKKKI KGIGLSI ILPFAIVANSYSTTLTVIANSVRYRLPMPIDICEKCEINQQCVSDICRCIMBEKE KODITRPKYDNIT QTKDSL ONGLDKHPFRYNSLNTHPSNDLKLODISHBESHEND LNIFEWSKRREPVO NNSSH CVSISTWEILTFPNRI SPWMXGIS LPAIL ILIYPLAI ELMNSQSIHQHLS ISPI IVLOJIL ILYUKLEPKXYTONEMA I KYTCIFFSKVEL HI NFPIILI KIICLVFSVNGCGTI NYIFS SYTNIGDN II KSYYII IGSSFYIVTYLL T YYLCILIH HNCCNVNNKKNINYVDDEVIL NTYFDYVNLPINHISY"	
	repeat__region complement(677..706) /note="AT_rich"	
	repeat__region /rpt_type=tandem 957..976 /rpt_type=tandem /rpt_unit="(taa) n"	
	repeat__region 1032..1068 /rpt_type=tandem /rpt_unit="(taaaa) n"	
	repeat__region 1453..1478 /rpt_type=tandem /rpt_unit="(a) n"	
	repeat__region complement(1520..1556) /note="AT_rich"	
	repeat__region /rpt_type=tandem complement(1584..1612) /note="AT_rich"	
	repeat__region complement(1680..1717) /note="AT_rich"	
	repeat__region /rpt_type=tandem complement(1737..1766) /note="AT_rich"	
	repeat__region complement(2814..2846) /note="AT_rich"	
	repeat__region /rpt_type=tandem complement(2936..3100) /rpt_type=tandem /rpt_unit="(tatcaa) n"	
	repeat__region 3015..3046 /rpt_type=tandem /rpt_unit="(taa) n"	
	repeat__region complement(3129..3184) /note="AT_rich"	
	repeat__region complement(3356..3496) /note="AT_rich"	
	repeat__region complement(3558..3651) /note="AT_rich"	
	repeat__region complement(3785..>4738 /rpt_type=tandem /locus_tag="PF11_0335"	

```

mRNA      <3785..>.4738  
CDS        /locus_tag="PF11_0335"  
           3785..4738  
           /locus_tag="PF11_0335"  
           /codon_start=1  
           /product="hypothetical protein"  
           /protein_id="AAW35919.1"  
           /db_xref="GI:23496261"  
           /translation="MNLLKRRLNNI FYGKRYKYDTOTNKENNNIKTNKKYVKHILQKL  
           TPMLCFILKYITTFIFELISNLHHVYENKKYLHKLKNNTLLSYKSQYUHLNQ  
           SYIKHNDEPNFLKVDKDKNEKTEKKSIQKNNNIQNLNNTHFKPKGYGGODKN  
           VYIHVIWNNTLYIVLTFTFIFELLPSFSYPKHKKENI FINSNPLYNINIDIOKI  
           RKPKVFLTRRFFFEELLPVOYTMBPYITWHINKHLKTSFILOSLLKKEVENCENIID  
           RKPFLYNQAQESYSLIGLICGSYLEHFOTKKKKRIDKNE"  
repeat__region complement(383). .3891)  
           /note="AT_rich"  
           /rpt_type=candem  
           complement(4077..4130)  
           /note="AT_rich"  
           /rpt_type=candem  
           complement(4179..4217)  
           /note="AT_rich"  
           /rpt_type=candem  
           complement(4326..4333)  
           /note="AT_rich"  
           /rpt_type=candem  
           4720..4757  
           /rpt_type=candem  
           /rpt_unit="(taaa)n"  
           complement(4750..4846)  
           /rpt_type=candem  
           /rpt_unit="(ta)n"  
           4762..4794  
           /rpt_type=candem  
           complement(4874..4917)  
           /note="AT_rich"  
           /rpt_type=candem  
           complement(<4910..>.6244)  
           /locus_tag="PF11_0336"  
           complement(join(<4910..5940,6115..>.6244))  
           /locus_tag="PF11_0336"  
           complement(join(<4910..5940,6115..6244))  
           /locus_tag="PF11_0336"  
           /codon_start=1  
           /product="hypothetical protein, conserved"  
           /protein_id="AAW35920.1"  
           /db_xref="GI:23496262"  
           /translation="MANRTDASAIKNGSNPOYLISNITSKIYDSPYMKCKPALT  
           ESIIQALNILEVGGTGGNRKPRFLCLIKLIQIQPDQDIYEYIKMDPVYLRLL  
           GIPYLRILTGLSELEVYNNLEPILPDYRKIRMRLONGTEPKLYMDFVNCCIILNNLDV  
           DPTLTKGVQLENNILLQRKLEYKALLNISDNELFNKNNEDTHANKKEINKNKE  
           YKORMLERHKHKKINKYSSEDDHNYHKKKEGYDKKKGHNKTYNHKYNNKQDDERRRH  
           SDXSDDMRKRRKKREKRTSKVDHKKKKKKRNYSDDSXSFTSSSEYNNKKCEHKK  
           KKKKLFPLKKOKKKXNHSSSOSNHSYDENKKDELSTDKMEIRKELGMKPLK"  
repeat__region complement(5025..5094)  
           /rpt_type=candem  
           /rpt_unit="(a)n"  
           complement(5145..5218)  
           /rpt_type=tandem  
           /rpt_unit="(a)n"  
           complement(5368..5388)  
           /note="AT_rich"  
           /rpt_type=candem  
           5945..5983  
           /rpt_type=candem  
           /rpt_unit="(a)n"  
           5985..6016  
           /rpt_type=candem  
           /rpt_unit="(tatata)n"  
           complement(6046..6113)  
           /note="AT_rich"  
           /rpt_type=candem  
           complement(6251..6290)

```



```

/note="AT rich"
/rpt_type=tandem
repeat_region
6419..6479
/rpt_type=tandem
repeat_region
complement(6419..6479)
/rpt_type=tandem
/rpt_unit="(a)n"
complement(6497..6686)
/note="AT rich"

```

Query Match 12.4%; Score 41.8; DB 3; Length 252420;
 Best Local Similarity 50.2%; Pred. No. 2.1;
 Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

```

QY 14 TTGATAGCGCGTTTCTTACTACACTCAATAGTGAACAGATGATGGTTAGTGA 73
DB 77046 TTATATTGATGATGATATATAGAAAATGCAATTAATAAATAGCAAGATTCATTTAA 77105
QY 74 CTGTTATATAAGAGATATAAGATATCATCATCATTTGAGCAATTAAGAGGAGGA 133
DB 77106 ATCTCTTTTATTAAGAAAATCAAAATGAGAGATTTCTTACAAAAGTTGTAATAGTA 77165
QY 134 GATTACGAAACAGTGTGCTTACAGTGAAGAAAACAATTAAAGTACCCCTCC 193
DB 77166 CAAGTAAATTAATTAACGTGCAACAAAGGGAACAAAAGATGTTATCAACAGG 77225
QY 194 TTGACAGATCAATGCCACGTTGA 218
DB 77226 TACATTAACAAAACAAAATTTGA 77250

```

RESULT 4
 AKI16155/c
 LOCUS AKI16155 1663 bp mRNA linear INV 30-Nov-2002
 DEFINITION Ciona intestinalis cDNA, clone:citb012j08, full insert sequence.
 ACCESSION AKI16155
 VERSION AKI16155.1 GI:23588595
 KEYWORDS FLI_CDNA.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.

```

REFERENCE
AUTHORS 1 Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawaashima,T.,
Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasaki,K.,
Nakayama,A., Ishikawa,H., Inaba,K. and Satoh,N.
TITLE A cDNA resource from the basal chordate Ciona intestinalis
JOURNAL Genesis 33 (4), 153-154 (2002)
MEDLINE 22191024
PUBMED 12203911
REFERENCE 2 (bases 1 to 1663)
AUTHORS Satou,Y. and Satoh,N.
TITLE Direct Submision
JOURNAL Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail:satoh@sci.kyoto-u.ac.jp, Tel:81-75-753-4095,
Fax:81-75-705-1113)
COMMENT Ciona intestinalis cDNA Project (URL:
http://ghost.zool.kyoto-u.ac.jp/index1.html).
FEATURES
source 1..1663
Location/Qualifiers
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb012j08"

```

Query Match 12.3%; Score 41.4; DB 3; Length 1663;
 Best Local Similarity 46.4%; Pred. No. 3.9;
 Matches 15; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

```

QY 35 ACNACTCAATTAAGATGAACAGAAATGAGCGTTAGTGAAGTCTTTTAAAGAAAGTAAAT 94
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 1602 ATTAAGTATTCACGAAAGACATATGTGTCCAAACGCTCAACAAATTAAGTAAAT 1543
QY 95 AAAGATCTATCATCATTTTGGAGCAATTAAGGAGGAGAGATTACGAAACAGTGTGCTT 154
DB 1542 GACAGTTCCAATACCTTTCTGTATAAATAATGACCCGTCTTCATTTATCATTTGACCA 1483
QY 155 ACAAGTGAACAACTTAACCTTAAGTAAAGTACCCCTCCCTGACAGATCAATGCCACAG 214
DB 1482 CGAAATGAAATAGCTTTAAATAGTAAACATTAATAAATGTTCACCAATGCAACCCCTG 1423
QY 215 TTGAGCTTATGACGACGACATCATCATGTAATGCTTCTGATTAAGCTGTTCATTA 274
DB 1422 CTAAAGCATCTGCTATGCTGATTAATTTCTGTGTCTTCTGACAACTAGTTGCAAC 1363
QY 275 TTCTCTTGCAGAACTCTGCTACTTACCAAGAACTGCTGCTACAGACAGAA 325
DB 1362 AGCTTCTGTAAATCTCCACACTTGATTCAGTTGCTGCTGCTGCTTAA 1312

```

RESULT 5
 OSJN00175/c
 LOCUS OSJN00175 99855 bp DNA linear PLN 10-FEB-2004
 DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBa0064D20,
 complete sequence.
 ACCESSION AL662973 GI:32489560
 VERSION AL662973.3 GI:32489560
 KEYWORDS HTG.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriaristidae; Oryzeae; Oryza.

```

REFERENCE
AUTHORS 1 Feng,Q., Zhang,Y., Hao,P., Wang,S., Fu,G., Huang,Y., Li,Y., Zhu,J.,
Liu,Y., Hu,X., Jia,P., Zhang,Y., Zhao,Q., Ying,K., Yu,S., Tang,Y.,
Weng,Q., Zhang,L., Lu,Y., Mu,J., Lu,Y., Zhang,L.S., Yu,Z., Fan,D.,
Liu,X., Lu,T., Li,C., Wu,Y., Sun,T., Lei,H., Li,T., Hu,H., Guan,J.,
Wu,M., Zhang,R., Zhou,B., Chen,Z., Chen,L., Jin,Z., Wang,R.,
Yin,H., Cai,Z., Ren,S., Lv,G., Gu,W., Zhu,G., Tu,Y., Jia,Q.,
Zhang,Y., Chen,J., Kang,H., Chen,X., Shao,C., Sun,Y., Jia,Q.,
Zhang,X., Zhang,W., Wang,L., Ding,C., Sheng,H., Gu,J., Chen,S.,
Ni,L., Zhu,F., Chen,W., Lan,L., Lai,Y., Cheng,Z., Gu,M., Jiang,J.,
Li,J., Hong,G., Xue,Y. and Han,B.
TITLE Sequence and analysis of rice chromosome 4
JOURNAL Nature 420 (6913), 316-320 (2002)
MEDLINE 12237377
PUBMED 12447439
REFERENCE 2
AUTHORS Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q., Hu,X.,
Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L., Weng,Q.J., Zhang,L.,
Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Lu,Y., Li,C.,
Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Ying,K., Zhou,B.,
Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Guan,J.P., Fu,G.,
Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F., Tu,Y.F.,
Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H., Chen,X.Y.,
Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W., Wang,L.J.,
Ding,C.W., Sheng,H.H., Gu,J.L., Chen,S.T., Ni,L., Zhu,F.H. and
Hong,G.F.
TITLE Direct Submision
JOURNAL Submitted (27-DEC-2001) Han Bn, National Center for Gene Research,
Chinese Academy of Sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
REMARK Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNBa0064D20.
COMMENT On Jul 9, 2003 this sequence version replaced gi:21912515.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),

```

genescan (<http://CCR-001.mit.edu/GENSCAN.html>), GeneMarkHM (<http://genemark.biology.gatech.edu/GeneMark/>), tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>), searches of the complete sequence against NCBI non-redundant protein database (<http://ncbi.nlm.nih.gov/blast/db>) and the EST database at NCBI. Location/Qualifiers
1..39855

```
source
1. .9985
/organism="Oryza sativa (japonica cultivar-group)"
```

```
complement(join(3866. .5543,5638. .6485,6573. .6899))
/gene="OSJNBa0064D20.1"
```

complement(10918,.12816)
/gene="OSJNBa0064D20.2"

```
25802. .25967,26575. .27420))
/gene="OSJNBa0064D20.3"
```

QAECDADVDALDSQWVPRILSLORLEDKVTFVQEBDELSMFYNLKRACCCSIASSA
NVGKAMEEHEHNDEHGDRRSLAAPSSPAAGLRAAPPIVSTRGREPLHGEAROR

```
complement join(29612..29827,29918..30088,30550..30590,
31036..31096,31762..32646))
```

QRAEMERIKISGUEGSDQVIFRCHQKNAUBIENBSECCURFQVUMWGF MWUNNO
TILITKI "
Comment (37600 41010)

gene complement (45448, .46707) N"

gene
KAT5RFFQSUNLALG
47424 . 50668
/cens= WGS.TMB.0064D30 711

```
./protein_id="CAE04224.2"
```

```
/db_xref="GI:38346577"
/db_xref="UniProt/TREMBL:Q7X729"

Query Match      12.3%; Score 41.4; DB 8; Length 9985;
Best Local Similarity 51.3%; Pred. No. 2.8;
Matches 96; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY      37 AACCTAAATAGATGACGAATGATGGTTGTGACGCTTTATTAAGAGAGTAAATA 96
DB      83615 AATATTGTAACCAACCAACTAGTCATGCTTTTCTTCAATTAAGAAATGATGCTGT 83556
QY      97 AATATCATCATCATTTGAGGCAATAGGAGGAGAGATTCGAAACAGTGTGCTTAC 156
DB      83555 ATTAATATCTGATATTGTGATATTTTAAGATTAATTAACGATGATTAAGTCTTTTA 83496
QY      157 AAGTGAACAACAGTTAACTAAAGTACCCCTCTTCAAGATCAATGCCACAGTT 216
DB      83495 AAGTGAACAAGTTGTTCTACAAAGATATATCTCCCTCATTCCTTAATGTTTGACACCGTT 83436
QY      217 GAGCTTT 223
DB      83435 GACTTTT 83429

RESULT 6
OSJN00292/c      121130 bp      DNA      linear      PLN 10-FEB-2004
LOCUS      Oryza sativa genomic DNA, chromosome 4, BAC clone: P0076017,
DEFINITION      complete sequence.
ACCESSION      BX548156
VERSION      BX548156.1 GI:32479667
KEYWORDS      HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM      Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE      1
AUTHORS      Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q., Hu,X.,
              Liu,Y.L., Mu,D., Yu,Z., Chen,L., Fan,D.L., Weng,Q.J., Zhang,L.,
              Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Lu,Y., Li,C.,
              Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Ying,K., Zhou,B.,
              Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Guan,J.P., Fu,G.,
              Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F., Tu,Y.X.,
              Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H., Chen,X.Y.,
              Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W., Wang,L.J.,
              Ding,C.W., Sheng,H.H., Gu,J.L., Chen,S.T., Ni,L., Zhu,F.H. and
              Hong,G.F.
              Direct Submision
              Submitted (08-JUL-2003) Han Bin, National Center for Gene Research,
              Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
              CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
              bhan@ncgr.ac.cn
              Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
              clone: P0076017.
              Web site: http://www.ncgr.ac.cn
              ----- Summary Statistics
              Assembly program: phrap
              -----
              This is a complete sequence.
              Genes were identified by a combination of several methods: Gene
              prediction programs including Fgenesh (http://www.softberry.com/),
              genscan (http://CCR-081.mit.edu/GENSCAN.html), GenemarkHM
              (http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean
              Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
              complete sequence against NCBI non-redundant protein database (nr)
              (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
              Location/Qualifiers
              1..121130
              /organism="Oryza sativa (japonica cultivar-group)"
              /mol_type="genomic DNA"
              /cultivar="Nipponbare"
              /sub_species="japonica"

gene
CDS
/db_xref="taxon:39947"
/chromosome="4"
/clone_1fb="P0076017"
/clone_1lb="RGF-05JNP"
/complement(4269..7724)
/gene="P0076017.1"
complement(join(4269..4451,4540..4695,5213..5330,
5458..6125,7695..7724))
/gene="P0076017.1"
/codon_start=1
/protein_id="CAE02503.1"
/db_xref="UniProt/TREMBL:Q7X902"
/translation="MASSFRKKTGMLIHPRPPRGCAVQTPSPPATDRLSALPDAL
LHATIIISLKGQVQTVLSKRWRLHMRVSPCLDIDQREFAASENWAISRDLKFE
FDADVLAVYRCSPAKLDTFLRLICDRYHSLSDSDIDWIRGLKSPREPHLPYR
YDSYLLENHKLGSNYGLTKLITVNSLHCEMBEHTTCTTLEVLNRCGLSYLOEI
THPKLKNLIVGHPVAMPISIVRASIRPSVFPCKQNFANLVFNVTLSLELSGPREMVE
FNQVFNPAFKNIRTLILDRCRPSYNNELRHILQNSPNLEKLTVHCCKRSKSLERK
SQHRKQVNCRKLTSTETIYKQIDIVRELVDLILDSGHLPRNTALTYKI"
11169..13339
/gene="P0076017.2"
join(11169..11467,12805..13339)
/gene="P0076017.2"
/codon_start=1
/protein_id="CAE02504.1"
/db_xref="UniProt/TREMBL:Q7X801"
/translation="MAPIPSKLSQGRPIAPVWSHLVYARNPPOESVLICRGTVPH
RHNPIASPHNPFYFOTHRHMSGFGSGNSPPKSTNNVCIGICICAPORARISRI
GMSYTEBGRRRPATTTFGRRORRAYGEORRRSGROLATGRLLALLPRAKERGSEEG
REMEYRAEIVTSGRRRSELSGLGPEKGKETAAPVHQLLARTPASPAAVEEGG
EDGSKAAKARWRNERNERBDGARFNRQCFRKEPTLAVKESBLAVIDACOSODG
R"
complement(16366..20771)
/gene="P0076017.3"
complement(join(16366..16494,18059..18151,18270..18367,
18906..18969,19067..19237,19926..20771))
/gene="P0076017.3"
/codon_start=1
/protein_id="CAE02505.1"
/db_xref="UniProt/TREMBL:Q7X877"
/translation="MEPAKARARTGGGSAVDRLSLPDLALHAWKSLPARQVQ
TCVLSKRWHLHMRVSPSLNDSREFLPIYDMQKQDEPTTLMFNHAPTDAFSIR
ADVAAGKGRHVDWIRGCIKCPVLIVAVTVSRVLPDLAAGSCGLGHLISYV
ALHSGPAROVRDSCVRLRCLELRCITKFSHESSTNRYLIVEDSMGSDSLAISAPR
IASRLIVAFLEFHYNGVSLNGANSIVASVAVSGRTSPBEAMULCGLFVSSELEI
KIQDLATLHKEFKDSQSPFNIRTLISLNCFOABGDLDKRFALGKLEMCNLEKUT
LOHCWFLGSGTQRAKRTKRTSGLCLENSSETLGDADVADVAALDSQVPRPILSLORLED
KVTFOEBDELSMFPNLRACCSIASANVVISRLKRGDEGDTLDSVPDPHENT
SQCPLKXIEIKYKLI"
complement(122963..25997)
/gene="P0076017.4"
complement(join(122963..23178,23269..23439,23901..23941,
24387..24447,25113..25997))
/gene="P0076017.4"
/codon_start=1
/protein_id="CAE02506.1"
/db_xref="UniProt/TREMBL:Q7X6F3"
/translation="MDRLSALPEGLHVMSPLPARONAQTCVLSKRWHLHMRVSPSL
NLDIIEFNNLEKEDDDDETGWKRPDIANLMPHNAFTLDTFTVTCGAAVOAAVGN
RSRRHNRNDVDPVAVANRSRDVNRIRGGIKYCPVLDVAVASGSCRPPDGS
GSCFRLRLRLSRSCALDSGPAROVRDSCVRLRCLELRCHEFHSIHSSSTNLEKYE
CIGCGLSLAISAPLALSLCIDISGAYKNGVSLNTIIVASVSTLNPVNSIPBEAM
LILCGFVNTNLELENIHAKLGGVZAKLSAVSAGSGRERKPTGCCVNEKVLHIE
KPFKFSFNNLRKTLISLDLDFQMGDLKRFALGKLEKCPVLEKTLLOHCWFLSGST
QRAEMETRTSGLCQSODQVTFHCKDILKIDSELCDHQLFOLMNGFWKDLKA
TIIITKI"
complement(31159..35270)
/gene="P0076017.5"
complement(join(31159..31263,32052..32222,32336..32486,
```


QY 257 GATTAAGCCCTGTTCAATATTCCTTTGCAAGCTCTGACTACTACCAAGAGTCTGCTAC 316
DB 114604 GGTATACCTTTTAAATATTCATTTCCCAAGCTTACTTCTTAAAAAGAGTCAATTTT 114663
QY 317 AGACAGAAAGATGGCATT 334
DB 114664 AGAAAGAAAAAGGTATT 114681

RESULT 10
AC026155 174294 bp DNA linear HTG 26-JUN-2001
AC026155/c
LOCUS
DEFINITION Homo sapiens chromosome 12 clone RP11-180F6, WORKING DRAFT
SEQUENCE 12 unordered pieces.
AC026155
AC026155.23 GI:14547355
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrook,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barbieri,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonini,D., Bouck,J.,
Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Diaper,H.,
Dugan-Rocha,S., Durbin,K.J., Earmhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Homel,P., Howard,S., Huber,J., Hulvik,S., Hume,J.,
Juckson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korval,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,W.,
Lousesed,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,G., Nguyen,N., Nickerson,E., Nwokemko,S.,
Ogih,M., Okunolu,G., Oregunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruitz,S., Savary,G., Scherer,S., Scott,G., Shubkan,H., Shooshari,N.,
Sisson,I., Sodergren,B., Sonalike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmari,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,A., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleciyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 174294)
Worley,K.C.
Direct Submission
Submitted (21-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced gi:13811703.
Center: Baylor College of Medicine

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project information
Center project name: HALP
Center clone name: RP11-180F6
Sequencing vector: Plasmid; M7789
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 176003 bases at least Q40
Consensus quality: 179810 bases at least Q30
Consensus quality: 181603 bases at least Q20
Estimated insert size: 178132; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 8.6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: this is a "working draft" sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 35024: contig of 35024 bp in length
* 35025 35124: gap of unknown length
* 35125 56161: contig of 21037 bp in length
* 56162 56261: gap of unknown length
* 56262 77352: contig of 21091 bp in length
* 77353 77452: gap of unknown length
* 77453 103958: contig of 26506 bp in length
* 103959 104058: gap of unknown length
* 104059 119954: contig of 15896 bp in length
* 119955 120054: gap of unknown length
* 120055 130550: contig of 10496 bp in length
* 130551 130650: gap of unknown length
* 130651 142703: contig of 12053 bp in length
* 142704 142803: gap of unknown length
* 142804 149562: contig of 6759 bp in length
* 149563 149662: gap of unknown length
* 149663 157532: contig of 7870 bp in length
* 157533 157632: gap of unknown length
* 157633 168762: contig of 11130 bp in length
* 168763 168862: gap of unknown length
* 168863 171961: contig of 3099 bp in length
* 171962 172061: gap of unknown length
* 172062 174294: contig of 2233 bp in length.

FEATURES
source
1..174294
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-180F6"

ORIGIN
Query Match 12.0%; Score 40.4; DB 2; Length 174294;
Best Local Similarity 55.8%; Pred No. 4.9; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 197 ACAGATCAATGCCACAGTTGAGCTTTAGCCAGCCACATCATGTAATTGCTTCT 256
DB 117449 AAAAGTATGATGCCCTCACTAGGTTTGAACAGTATTTTACATGATTAAGATTACCT 117390
QY 257 GATTAAGCCCTGTTCAATATTCCTTTGCAAGCTCTGCTACTACTACCAAGAGTCTGCTAC 316
DB 117389 GGTATACCTTTTAAATATTCATTTCCCAAGCTTACTTCTTAAAAAGAGTCAATTTT 117350
QY 317 AGACAGAAAGATGGCATT 334


```

repeat_region      14235..14272
                    /rpt_family="(CA)n"
repeat_region      14521..14558
                    /rpt_family="(TTA)n"
repeat_region      14935..14963
                    /rpt_family="AT_rich"
repeat_region      15589..15629
                    /rpt_family="(TTTTC)n"
repeat_region      complement(15632..15765)
                    /rpt_family="B1_MM"
repeat_region      15919..16063
                    /rpt_family="B1_MM"
repeat_region      16068..16090
                    /rpt_family="(CAAC)n"
repeat_region      16143..16166
                    /rpt_family="(CAA)n"
repeat_region      complement(17349..17514)
                    /rpt_family="RSINE1"
repeat_region      18604..18631
                    /rpt_family="(A)n"
repeat_region      18696..18735
                    /rpt_family="(A)n"
repeat_region      19686..19753
                    /rpt_family="CT_rich"
repeat_region      22500..22598
                    /rpt_family="(TTTTC)n"
repeat_region      22891..22945
                    /rpt_family="CT_rich"
repeat_region      complement(24470..24595)
                    /rpt_family="RSINE1"
repeat_region      25058..25105
                    /rpt_family="(TATTA)n"
repeat_region      26176..26465
                    /rpt_family="B4A"
repeat_region      26473..26541
                    /rpt_family="(CA)n"
repeat_region      26917..26939
                    /rpt_family="(CAAA)n"
repeat_region      28766..28877
                    /rpt_family="PBID9"
repeat_region      28878..28909
                    /rpt_family="(CAA)n"
repeat_region      29243..29455
                    /rpt_family="B4"
repeat_region      30293..30336
                    /rpt_family="AT_rich"
repeat_region      complement(30557..31009)
                    /rpt_family="L1"
repeat_region      complement(31159..31213)
                    /rpt_family="L1"
repeat_region      31219..32472
                    /rpt_family="Lx4"
repeat_region      complement(33069..33221)
                    /rpt_family="L1_MM"
repeat_region      33438..33470
                    /rpt_family="(CAA)n"
repeat_region      complement(36123..36255)
                    /rpt_family="B1F"
repeat_region      37393..37422
                    /rpt_family="AT_rich"
repeat_region      40393..40422
                    /rpt_family="(TA)n"
repeat_region      40692..40712
                    /rpt_family="AT_rich"
repeat_region      41263..41297
                    /rpt_family="AT_rich"
repeat_region      42833..42861
                    /rpt_family="AT_rich"
repeat_region      43372..43402
                    /rpt_family="AT_rich"
repeat_region      complement(44537..44680)
                    /rpt_family="B1_MM"
repeat_region      complement(44781..44978)

```

```

repeat_region      44993..45026
                    /rpt_family="AT_rich"
repeat_region      45864..45939
                    /rpt_family="CT_rich"
repeat_region      45948..45982
                    /rpt_family="(CA)n"
repeat_region      complement(47163..47222)
                    /rpt_family="SS"
repeat_region      47927..48187
                    /rpt_family="Lx8"
repeat_region      48211..48294
                    /rpt_family="Lx8"
repeat_region      48808..48990
                    /rpt_family="B3"
repeat_region      complement(50184..50368)
                    /rpt_family="B3"
repeat_region      50559..50602
                    /rpt_family="(TAA)n"

```

Query Match 12.0%; Score 40.2; DB 10; Length 201783;
 Best local Similarity 48.9%; Pred.No.5.4; Indels 0; Gaps 0;
 Matches 108; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

```

QY      26 TTTTCTACTACACTCAATTAAGTAAAGATGATGCTTACTGCTTTTAAAG 85
DB      87178 TTTTATAGCTAAATTTGAAAATTTGCTTCTCTGTATTTGTTATTAATAT 87119
QY      86 AAGATTAATAAGATCATCTATCTATTTGAGCAATTAAGGAGAGATTCAGCAAC 145
DB      87118 CAGACATGTTTCTTTCAGCTCAGTGAAGGTTATTTGAGGAGAGACCACTG 87059
QY      146 AGTGTGCTTACAGTGAAGAAACAAGTTAAAGTGAAGCCCTCTGCAAGATCA 205
DB      87058 AGTGTGAGACAGTGCAGCACACATGACCTTGCGCAAGGCTTTCATTAAAGACA 86999
QY      206 ATGCCACAGTTGAGCTTTAGCCAGCCACATCATCTATTA 246
DB      86998 GGGAGAGGAGGAGTGAAGGGCGTGCACATCTTCA 86958

```

```

RESULT 13
CR318618/c      177655 bp      DNA      linear      HTG 01-MAR-2004
LOCUS
DEFINITION
unordered pieces.
ACCESSION
CR318618
VERSION
CR318618.1 GI:44864762
KEYWORDS
HTG; HTGS PHASE1.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 177655)
McLay, K.
REFERENCE
Direct Submission
Submitted (27-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: ZC3917
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 175828 bases at least Q40
Consensus quality: 176161 bases at least Q30
Consensus quality: 176379 bases at least Q20

```

COMMENT

Insert size: 176755; sum-of-contigs
 Insert size: 188044; 8.4% error; agarose-fp
 Quality coverage: 9.57x in Q20 bases; sum-of-contigs Quality
 coverage: 9.27x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 57111: contig of 57111 bp in length
* 57112 57211: gap of 100 bp
* 57212 73074: contig of 15863 bp in length
* 73075 73174: gap of 100 bp
* 73175 80257: contig of 7082 bp in length
* 80257 80357: gap of 100 bp
* 80357 89661: contig of 9304 bp in length
* 89661 89761: gap of 100 bp
* 89761 98899: contig of 9139 bp in length
* 98899 99000: gap of 100 bp
* 99000 105065: contig of 6066 bp in length
* 105065 105165: gap of 100 bp
* 105165 130251: contig of 25086 bp in length
* 130251 130352: gap of 100 bp
* 130352 166939: contig of 36588 bp in length
* 166939 167040: gap of 100 bp
* 167040 172134: contig of 5095 bp in length
* 172134 172235: gap of 100 bp
* 172235 177655: contig of 5421 bp in length.

```

FEATURES

```

source
  1..177655
  /organism="Danio rerio"
  /mol_type="genomic DNA"
  /db_xref="taxon:7955"
  /clone_id="CH211-3917"
  /clone_lib="CHORI-211"
  1..57111
  /note="assembly fragment:01750"
  fragment_chain:1
  clone_end:SP6
  vector_side:left"
  57212..73074
  /note="assembly fragment:00526"
  fragment_chain:1"
  73175..80256
  /note="assembly fragment:00319"
  fragment_chain:1"
  80357..89660
  /note="assembly fragment:00418"
  fragment_chain:2"
  89761..98899
  /note="assembly fragment:00153"
  fragment_chain:2"
  99000..105065
  /note="assembly fragment:00044"
  fragment_chain:2"
  105166..130251
  /note="assembly fragment:00847"
  fragment_chain:2"
  130352..166939
  /note="assembly fragment:01200"
  fragment_chain:2"
  167040..172134
  /note="assembly fragment:00235"
  fragment_chain:2"
  172235..177655
  /note="assembly fragment:00087"
  fragment_chain:2"

```

ORIGIN

Query Match 11.8%; Score 39.8; DB 2; Length 177655;
 Best Local Similarity 52.8%; Pred. No. 6.9;
 Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

Qy 1 ATTATTATTTTCATTGATAGCCGCTTTTACTACAACTCAATAGATGACAGATG 60
Db 116279 ATACATTTTATTCACAGACACTCAATTTTAAATATATATATATGATTAATG 116220
Qy 61 AATGGGTTAGTCAGCTGTTTAAAGAGATTAATAGATCAATCATCTTTGAGCAA 120
Db 116219 AATTTAAAGTGAATATATATATAGAAATGTCACATTTCTATATTTATTAAGTCA 116160
Qy 121 TAAGGAGGAGAGATTTCAGCAACAGTGTGCTTACAGTGA 163
Db 116159 CATACAGTGTCTCAGCATATACAGACACCCCTTACAAATGTA 116117

```

RESULT 14

BX537337

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITL

JOURNAL

COMMENT

Submitted (17-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 zfish-help@sanger.ac.uk
 On Apr 19, 2004 this sequence version replaced gi:46194230.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMPEP, Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 Zebrafish pUC subclones occasionally display inconsistency over the
 length of mononucleotide A/T runs and conserved TA repeats. Where
 this is found the longest good quality representation will be
 submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat
 discovery system (Zhirong Bao and Sean Eddy, submitted), and those
 beginning 'dir' were identified by Rick Waterman (Stephen Johnson
 lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmark.shtml
 CH211-222E23 is from a CHORI-211 BAC library
 VECTOR: pTRABAC2.1.

FEATURES
source Location/Qualifiers
1. 217375
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-222E23"
/clone_1ib="CHORI-211"

ORIGIN

Query Match 11.8%; Score 39.8; DB 5; Length 217375;
Best Local Similarity 52.8%; Pred. No. 6.8;
Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

1 ATTATATTTCTTCAATAGCGCGTTTCTTACACACATCAATATGAAACAGATG 60
157782 ATTCATTTTATTCAGACACCTCAATTTTAAATATATATGAAATTTAAATG 157841

61 AATGGTTAGTGACTGTTTATTAAGAAGATTAATTAAGTATCATCATTTGAGCAA 120
157842 AATTTAAAGTGAATATATATATGAAATGTCACATTTCTATATTTATTAAGTCAA 157901

121 TAAGGAGGAGAGATTCAGCAACAGCTGCTTACAGTGA 163
157902 CATACAGTCTCAGCATATACAGACACCCCTTACAAATGTA 157944

RESULT 15
AC068698/c 162605 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 14 clone RP11-555P23 map 14, WORKING DRAFT
DEFINITION
ACCESSION AC068698
VERSION AC068698.2 GI:8565695
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 162605)
Birren,B., Linton,L., Nuebaum,C. and Lander,E.
Homo sapiens chromosome 14, clone RP11-555P23
Unpublished
2 (bases 1 to 162605)
Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barina,N., Baetien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testfay,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 162605)
Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barina,N., Baetien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testfay,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

COMMENT
JOURNAL
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 16, 2000 this sequence version replaced gi:7712243.
ALL repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 555_P_23
Center clone name: 555_P_23
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Chemistry: Phrap; version 0.960731
Consensus quality: 154906 bases at least Q40
Consensus quality: 159118 bases at least Q30
Consensus quality: 160511 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 161405; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs

Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Landers,T., Lehoczy,J.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testfay,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 16, 2000 this sequence version replaced gi:7712243.
ALL repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 555_P_23
Center clone name: 555_P_23
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Chemistry: Phrap; version 0.960731
Consensus quality: 154906 bases at least Q40
Consensus quality: 159118 bases at least Q30
Consensus quality: 160511 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 161405; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs

NOTES: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 464: contig of 464 bp in length
465 564: gap of 100 bp
565 2397: contig of 1833 bp in length
2398 2497: gap of 100 bp
2498 4504: contig of 2007 bp in length
4505 4604: gap of 100 bp
4605 10523: contig of 5919 bp in length
10524 10623: gap of 100 bp
10624 18758: contig of 8135 bp in length
18759 18858: gap of 100 bp
18859 27777: contig of 8919 bp in length
27778 27877: gap of 100 bp
27878 40462: contig of 12585 bp in length
40463 40562: gap of 100 bp
40563 53834: contig of 13772 bp in length
53835 53934: gap of 100 bp
53935 69669: contig of 15935 bp in length
69670 69870: gap of 100 bp
69871 87514: contig of 17545 bp in length
87515 87615: gap of 100 bp
87616 103384: contig of 15770 bp in length
103385 103484: gap of 100 bp

* 103485 129543: contig of 26059 bp in length
* 129544 129643: gap of 100 bp
* 129644 162605: contig of 32962 bp in length.

FEATURES

source

1.162605
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/map="14"
/clone="RP11-555P23"
/clone_11b="RPC1-11 Human Male BAC"
1.464
/note="assembly_fragment"
clone_end:T7
vector_side:left
565..2397
/note="assembly_fragment"
2498..4504
/note="assembly_fragment"
4605..10523
/note="assembly_fragment"
clone_end:SP6
vector_side:right
10624..18758
/note="assembly_fragment"
18859..27777
/note="assembly_fragment"
27878..40462
/note="assembly_fragment"
40563..53834
/note="assembly_fragment"
53935..69869
/note="assembly_fragment"
69970..87514
/note="assembly_fragment"
87615..103384
/note="assembly_fragment"
103485..129543
/note="assembly_fragment"
129644..162605
/note="assembly_fragment"

ORIGIN

Query Match 11.8%; Score 39.6; DB 2; Length 162605;
Best Local Similarity 51.7%; Pred. No. 7.9;
Matches 90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 18 ATAGCCGGTTTTCACACACTCAATATAGATGACAGAAATGATGCGTTAGTACTGT 77
Db 104024 ACAGTCCCTTCCCACTACAGGTCATACAAAGAGAGGAAAAGCTTGAGTTGGATA 103965
QY 78 TTATTAAGAAGAGTAATTAAGATATCATCTTTCAGGCAATTAAGGAGAGAGAT 137
Db 103964 TGAATTTGAAGTTTGTATACGAATTAATCTTACTTTAGTAAGTGAAGTCTTA 103905
QY 138 CAGCAAACTGCTTACAAAGTGAACAACTTAATTAAGTGAACCCCT 191
Db 103904 CCATTAATAGGCTATTATTAAGTGAAGATGAAAAAATGATGGGCTTTCCT 103851

Search completed: November 21, 2005, 07:00:45
Job time: 1592.72 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 07:00:52 ; Search time 818.717 Seconds
(without alignments)
5171.686 Million cell updates/sec

Title: US-10-099-663-1_COPY_1115_1626

Perfect score: 512
Sequence: 1 cagctgactctcatctcat.....gtctgctacagacagaag 512

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications NA:*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_NEW_PUB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
18:	/cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
19:	/cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
20:	/cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
21:	/cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
22:	/cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
23:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
24:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq2:*
25:	/cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
26:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
27:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
28:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	512	100.0	2381	17	US-10-099-663-1	Sequence 1, Appl1
2	326	63.7	336	17	US-10-099-663-2	Sequence 2, Appl1
3	43.6	8.5	413	20	US-10-674-124A-2159	Sequence 2159, Ap
4	42.8	8.4	558	13	US-09-925-065A-572321	Sequence 572321,
5	42.8	8.4	679	13	US-09-925-065A-875384	Sequence 875384,

C	6	42.8	8.4	679	13	US-09-925-065A-875385	Sequence 875385,
C	7	42.8	8.4	679	13	US-09-925-065A-875386	Sequence 875386,
C	8	42.4	8.3	5908	15	US-10-239-676-94	Sequence 94, Appl1
C	9	42.4	8.3	5908	16	US-10-240-453-106	Sequence 106, Appl
C	10	42.4	8.3	5908	18	US-10-221-613-176	Sequence 176, App
C	11	42.2	8.2	1238	13	US-09-925-065A-710652	Sequence 710652,
C	12	41.8	8.2	10048	18	US-10-221-613-211	Sequence 211, App
C	13	40.8	8.0	546	13	US-09-925-065A-121976	Sequence 121976,
C	14	40.6	7.9	7040	16	US-10-172-086-13	Sequence 13, Appl
C	15	40.6	7.9	7040	19	US-10-221-714A-161	Sequence 161, Appl
C	16	40.6	7.9	7040	20	US-10-311-507-47	Sequence 47, Appl
C	17	40.6	7.9	7040	21	US-10-480-846-13	Sequence 173, Appl
C	18	40.6	7.9	7040	21	US-10-473-126-173	Sequence 173, App
C	19	40.6	7.9	7040	21	US-10-473-126-319	Sequence 319, App
C	20	40.6	7.9	14095	16	US-10-311-455-450	Sequence 450, App
C	21	40.4	7.9	470	13	US-09-925-065A-178215	Sequence 178215,
C	22	40.4	7.9	1816	13	US-09-925-065A-178216	Sequence 178216,
C	23	40.4	7.9	1816	13	US-09-925-065A-45365	Sequence 45365, A
C	24	40.2	7.9	6065	18	US-10-221-613-221	Sequence 221, App
C	25	39.4	7.7	5265	16	US-10-240-453-283	Sequence 293, App
C	26	39.2	7.7	9117	16	US-10-311-455-1042	Sequence 1042, App
C	27	39.2	7.6	2933	26	US-11-097-143-25165	Sequence 25165, A
C	28	39.2	7.6	5311	16	US-10-311-455-992	Sequence 992, App
C	29	39.2	7.6	6202	15	US-10-175-523-189	Sequence 189, App
C	30	39.2	7.6	6202	26	US-11-099-266-189	Sequence 189, App
C	31	39.2	7.6	6202	19	US-10-336-603A-35	Sequence 35, Appl
C	32	39.2	7.6	22052	26	US-11-097-143-28054	Sequence 28054, A
C	33	39.2	7.6	26097	26	US-11-097-143-2785	Sequence 2785, Ap
C	34	39.2	7.6	36194	26	US-11-097-143-28114	Sequence 28114, A
C	35	39.2	7.6	59667	26	US-11-097-143-20479	Sequence 20479, A
C	36	38.8	7.6	1024	21	US-10-425-115-97707	Sequence 97707, A
C	37	38.4	7.5	32392	22	US-10-706-633-27	Sequence 27, Appl
C	38	38.4	7.5	50000	22	US-10-706-635-23	Sequence 23, Appl
C	39	38.2	7.5	6072	16	US-10-311-455-4	Sequence 4, Appl
C	40	38.2	7.5	56737	9	US-09-782-378A-17	Sequence 17, Appl
C	41	38.2	7.5	684973	9	US-09-263-959-1	Sequence 1, Appl
C	42	38	7.4	6802	16	US-10-311-455-1204	Sequence 1204, Ap
C	43	38	7.4	640881	9	US-09-790-988-1	Sequence 1, Appl
C	44	37.8	7.4	3322	26	US-11-097-143-13510	Sequence 13510, A
C	45	37.8	7.4	17527	16	US-10-311-455-1406	Sequence 1406, Ap

ALIGNMENTS

RESULT 1
US-10-099-663-1
; Sequence 1, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenice, Inc
; TITLE OF INVENTION: Avian GUT-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(1626)
; NAME/KEY: exon
; LOCATION: (1627)..(1693)
; NAME/KEY: Intron
; LOCATION: (1694)..(2322)
; NAME/KEY: exon
; LOCATION: (2333)..(2381)
US-10-099-663-1
Query Match 100.0%; Score 512; DB 17; Length 2381;

Best Local Similarity 100.0%; Pred. No. 8.1e-124;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CAGTGATCTTCATCTCATGCTTATTATGTGGAGTAGAATGATTCACCAATTAGA 60
DB 1115 CAGTGATCTTCATCTCATGCTTATTATGTGGAGTAGAATGATTCACCAATTAGA 1174
QY 61 ATGGACAAGAGAGATTTGTTTATCTGTGGTAAATACGTTTCTCCAGTTGAT 120
DB 1175 ATGGACAAGAGAGATTTGTTTATCTGTGGTAAATACGTTTCTCCAGTTGAT 1234
QY 121 AAAGACCTCCACACAGATTAAGTCTTATGCAACAAGAAATGCAATGATCTCTT 180
DB 1235 AAAGACCTCCACACAGATTAAGTCTTATGCAACAAGAAATGCAATGATCTCTT 1294
QY 181 AGTCTATTATTATTTTCACTAGATAGCCGGTTTCTTACTACACTCAAAATTAAGTAAAC 240
DB 1295 AGTCTATTATTATTTTCACTAGATAGCCGGTTTCTTACTACACTCAAAATTAAGTAAAC 1354
QY 241 AGAATGAATGGGTAGTGACTGTTTATAAGAAAGATTAATTAAGATACATCATTTTG 300
DB 1355 AGAATGAATGGGTAGTGACTGTTTATAAGAAAGATTAATTAAGATACATCATTTTG 1414
QY 301 AGGCAATTAAGGAGGAGAGATTCAGCAAAACAGTGTGCTTACAGTGAAGAAACAAGTTAA 360
DB 1415 AGGCAATTAAGGAGGAGAGATTCAGCAAAACAGTGTGCTTACAGTGAAGAAACAAGTTAA 1474
QY 361 ACTAAAGTACCCCTCTCTTGCAAGATCATGCAACAGTGAAGTTAGCCAGCCACA 420
DB 1475 ACTAAAGTACCCCTCTCTTGCAAGATCATGCAACAGTGAAGTTAGCCAGCCACA 1534
QY 421 TCATCATGTAATTTGCTTCTTGATTAAGCCTGTTCAATTAATCTCTTGCAAAAGCTCTGC 480
DB 1535 TCATCATGTAATTTGCTTCTTGATTAAGCCTGTTCAATTAATCTCTTGCAAAAGCTCTGC 1594
QY 481 TACTTACAGAAAGTCTGCTTACAGACAGAAAG 512
DB 1595 TACTTACAGAAAGTCTGCTTACAGACAGAAAG 1626
```

```
RESULT 2
US-10-099-663-2
; Sequence 2, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Gut-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(336)
; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region
US-10-099-663-2
```

Query Match 63.7%; Score 326; DB 17; Length 336;
Best Local Similarity 100.0%; Pred. No. 2.6e-75;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 187 ATTATATTTTCTAGATAGCCGTTTCTTACTACAACTCAATTAAGATGAAGATG 246
DB 1 ATTATATTTTCTAGATAGCCGTTTCTTACTACAACTCAATTAAGATGAAGATG 60
QY 247 AATGGTTAGTGACTGTTTATAAGAAAGATTAATTAAGATCACTATCATCTTTGAGCAA 306
DB 61 AATGGTTAGTGACTGTTTATAAGAAAGATTAATTAAGATCACTATCATCTTTGAGCAA 120
```

```
QY 307 TAAGGAGGAGAGATTCAGCAAAACAGTGTGCTTACAAGTGGAAAACAAGTTAAACTAA 366
DB 121 TAAGGAGGAGAGATTCAGCAAAACAGTGTGCTTACAAGTGGAAAACAAGTTAAACTAA 180
QY 367 GTACCCCTCTCTTGAACAAGATCAATCCACAGTTGAGCTTTAGCCACCATCATCA 426
DB 181 GTACCCCTCTCTTGAACAAGATCAATCCACAGTTGAGCTTTAGCCACCATCATCA 240
QY 427 TGTAAATGCTTCTGATTAAGCCTGTTCAATTAATTTCTTTGCAAAAGCTGCTACTTA 486
DB 241 TGTAAATGCTTCTGATTAAGCCTGTTCAATTAATTTCTTTGCAAAAGCTGCTACTTA 300
QY 487 CCAGAAAGTCTGCTTACAGACAGAAAG 512
DB 301 CCAGAAAGTCTGCTTACAGACAGAAAG 326
```

```
RESULT 3
US-10-674-124A-2159/c
; Sequence 2159, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hideoshi
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: ORIN-003C1P
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 2159
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: AC012079.4_71311
; FEATURE:
; OTHER INFORMATION: Located on chromosome 1
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: sequence : 250648904
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 205277
US-10-674-124A-2159
```

Query Match 8.5%; Score 43.6; DB 20; Length 413;
Best Local Similarity 55.2%; Pred. No. 0.47;
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

```
QY 138 TATTAAGTCTTATGCAACAAGAAATGCAATTAATCTCTTACTGCTATTATTTT 197
DB 256 TATTAAGTCTTATGCAACAAGTCTTATTAATTAATTAATTAATTAATTAATTAAT 197
QY 198 CATTAAGTACCGGTTTCTTACTACACTCAATTAAGATGAAGAAAGATGAGTTAGT 257
DB 196 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 137
QY 258 GACTGTTTATAAGAAAGATTAATTAAGATCACTAT 291
```

Db 136 GACCTTATAGTAGAATAATGTAATCTAATTT 103

RESULT 4

US-09-925-065A-572321/c
; Sequence 572321, Application US/09925065A
; Publication No. US20050228172A9

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR FILING DATE: 2000-10-24

PRIOR FILING DATE: 2000-11-20

PRIOR FILING DATE: 2000-11-20

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 572321

LENGTH: 558

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-572321

Query Match

Best Local Similarity 8.4%; Score 42.8; DB 13; Length 558;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 187 ATTATATTTTTCATTAGATAGCGGTTTTTTTCTACAACTCAATAGTGAACGAATG 246

Db 346 ATTAACATTAAACAAATTAGTACCTTTATATCTTTTAAACAACGAACTTTAGAACT 287

Qy 247 AATGGTTAGTACCTTTTATTAAGAGTAATAAGTACTATCATCTATTGAGGCA 306

Db 286 TAAGTAAGTCCATCTATTAATAGAAATAGAACTATTAATAGAAATAGCA 227

Qy 307 TAAGGAGGAGAGATTTCAGCAACAGTGCTTCAACTGGAAAA 352

Db 226 TAAGAACTAATACTTAGAATAATTAATAAAAAATGTACAAA 181

Db 226 TAAGAACTAATACTTAGAATAATTAATAAAAAATGTACAAA 181

RESULT 5

US-09-925-065A-875384/c
; Sequence 875384, Application US/09925065A
; Publication No. US20050228172A9

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR FILING DATE: 2000-10-24

PRIOR FILING DATE: 2000-11-20

PRIOR FILING DATE: 2000-11-20

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 875384

LENGTH: 679

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-875384

Query Match

Best Local Similarity 8.4%; Score 42.8; DB 13; Length 679;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 187 ATTATATTTTTCATTAGATAGCGGTTTTTTTCTACAACTCAATAGTGAACGAATG 246

Db 515 ATTAACATTAAACAAATTAGTACCTTTATATCTTTTAAACAACGAACTTTAGAACT 456

Qy 247 AATGGTTAGTACCTTTTATTAAGAGTAATAAGTACTATCATCTATTGAGGCA 306

Db 455 TAAGTAAGTCCATCTATTAATAGAAATAGAACTATTAATAGAAATAGCA 396

Qy 307 TAAGGAGGAGAGATTTCAGCAACAGTGCTTCAACTGGAAAA 352

Db 395 TAAGAACTAATACTTAGAATAATTAATAAAAAATGTACAAA 350

Db 395 TAAGAACTAATACTTAGAATAATTAATAAAAAATGTACAAA 350

RESULT 6

US-09-925-065A-875385/c
; Sequence 875385, Application US/09925065A
; Publication No. US20050228172A9

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR FILING DATE: 2000-10-24

PRIOR FILING DATE: 2000-11-20

PRIOR FILING DATE: 2000-11-20

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 875385

LENGTH: 679

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-875385

Query Match

Best Local Similarity 8.4%; Score 42.8; DB 13; Length 679;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 187 ATTATATTTTTCATTAGATAGCGGTTTTTTTCTACAACTCAATAGTGAACGAATG 246

Db 515 ATTAACATTAAACAAATTAGTACCTTTATATCTTTTAAACAACGAACTTTAGAACT 456

Qy 247 AATGGTTAGTACCTTTTATTAAGAGTAATAAGTACTATCATCTATTGAGGCA 306

Db 455 TAAGTAAGTCCATCTATTAATAGAAATAGAACTATTAATAGAAATAGCA 396

Qy 307 TAAGGAGGAGAGATTTCAGCAACAGTGCTTCAACTGGAAAA 352

Db 395 TAAGAACTAATACTTAGAATAATTAATAAAAAATGTACAAA 350

Db 395 TAAGAACTAATACTTAGAATAATTAATAAAAAATGTACAAA 350

RESULT 7

US-09-925-065A-875386/c
; Sequence 875386, Application US/09925065A
; Publication No. US20050228172A9

GENERAL INFORMATION:

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 875386
LENGTH: 679
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-875386
```

```
Query Match
Best Local Similarity 53.6%; Pred. No. 0.94; Length 679;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```
QY 187 ATTATATTTTCTAGATAGCGGTTTCTTACACTCAATTAAGATGAACAGATG 246
    |||||
DB 515 ATTAACTTAACAAATTAAGCTTTATATCTTTTAAACAAGCACTTTAGAACT 456

QY 247 AATGGTGTGACCTTTTATAAGAGCTAATAAGATCATTCATTTGAGGCA 306
    |||||
DB 455 TAAGAGTAAGTCCATTAATTAAGATTAAGATTAAGATTAAGATTAAGATTA 396

QY 307 TAAGAGGAGAGATTCAGCAACAGTGTGCTTACAGTGGAAAA 352
    |||||
DB 395 TAAGAACACTTAATTAATTAAGATTAATTAATTAATTAATTAATTAATTA 350
```

```
RESULT 8
US-10-239-676-94/c
; Sequence 94, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10033529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 94
; LENGTH: 5908
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-94
```

```
Query Match 8.3%; Score 42.4; DB 15; Length 5908;
```

```
Best Local Similarity 54.5%; Pred. No. 3.1;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 138 TATTAAGTCCTATGCAACAAAGAAATGCAATTCCTTACTGCTATTTATTT 197
    |||||
DB 4703 TATACTCTTTTATTAATAAACAATATATCTATATCTAATAACCACTTTTAA 4644

QY 198 CATTAGATAGCGGTTTCTTACTACACTCAATTAAGATGAACAGATGAGTGGT 257
    |||||
DB 4643 TATTATATTTTATTTCACTACTCTTCTTAATAATTAATAATTAATTAATTA 4584

QY 258 GACTGTTTATAAAGAGATTAATAAGATCTATCA 293
    |||||
DB 4583 TCCCTTTATCCAAATTAATTAATCTATATAGCAACA 4548
```

```
RESULT 9
US-10-240-453-106/c
; Sequence 106, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 106
; LENGTH: 5908
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-453-106
```

```
Query Match
Best Local Similarity 54.5%; Pred. No. 3.1; Length 5908;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
```

```
QY 138 TATTAAGTCCTATGCAACAAAGAAATGCAATTCCTTACTGCTATTTATTT 197
    |||||
DB 4703 TATACTCTTTTATTAATAAACAATATATCTATATCTAATAACCACTTTTAA 4644

QY 198 CATTAGATAGCGGTTTCTTACTACACTCAATTAAGATGAACAGATGAGTGGT 257
    |||||
DB 4643 TATTATATTTTATTTCACTACTCTTCTTAATAATTAATAATTAATTAATTA 4584

QY 258 GACTGTTTATAAAGAGATTAATAAGATCTATCA 293
    |||||
DB 4583 TCCCTTTATCCAAATTAATTAATCTATATAGCAACA 4548
```

```
RESULT 10
US-10-221-613-176/c
; Sequence 176, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
```



```

; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 176
; LENGTH: 5908
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-176

Query Match
Best Local Similarity 8.3%; Score 42.4; DB 18; Length 5908;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 138 TATTAAGTCTCTATGCAACAAGAAATGTCATACCTCTCTAGTCTATTATTTT 197
DB 4703 TATTAAGTCTCTATGCAACAAGAAATGTCATACCTCTCTAGTCTATTATTTT 197
QY 198 CATTAAGTCTCTATGCAACAAGAAATGTCATACCTCTCTAGTCTATTATTTT 257
DB 4643 TATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 4584
QY 258 GACTGTTTATTAAGAAAGAGTAAATTAAGATATCA 293
DB 4583 TCCCTTTATCCAAATTAATTAATCTATATACGACA 4548

RESULT 11
US-09-925-065A-710652/c
; Sequence 710652, Application US/03925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 710652
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-710652
```

```

Query Match
Best Local Similarity 8.2%; Score 42.2; DB 13; Length 1238;
Matches 92; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 74 AGATTGTGTTTATCTGTTGGTAAATAGTTTCTCCAGTTGTATTAAGACCTCCCA 133
DB 986 ATATTGAGGTTTCTTGGCTGACATATCAATTTTATGAGTAAGTAAAGCATRCAA 927
QY 134 CCAGTATTAAGTCTATGCAACAAGAAATGTCATATCTCTTATGTCATTATTA 193
DB 926 ATCTCATGTAATGATATGATGATATGATGCTCTCAATGATGATTTTCTTATTT 867
QY 194 TTTTATGATAGCCGTTTCTTATCAATCTCAATTAAGTAAAGTAAAGTAA 248
DB 866 CTATATATTAATTAAGATATTTTGTGTAATTCGAAATCTAAATCAATTGAA 812

RESULT 12
US-10-221-613-211/c
; Sequence 211, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 211
; LENGTH: 10048
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3184, 3255, 3396, 6016, 8568)
US-10-221-613-211

Query Match
Best Local Similarity 46.4%; Score 41.8; DB 18; Length 10048;
Matches 136; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 100 ATACGTTTCTCCAGTTGTATTAAGACCTCCACAGATTAAGCTATGCAACAAG 159
DB 5471 ATATTATTTATCTTAAATCTAAATTCCTCTCTACAAAAAATTAATTAATCTAT 5412
QY 160 AAATGTCATATCTCTTATGTCATATTTATTTATTTATTTATTTATTTATTT 219
DB 5411 TCCAAATCCAAAAATCTATTTTAAACCTATTAATTCCTCTCAATTAATCTTCTAC 5352
QY 220 TCAACTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 279
DB 5351 AAATGTCATATCTCTTATGTCATATTTATTTATTTATTTATTTATTTATTT 5292
QY 280 TAAAGTATCATCATTTGAGCAATTAAGGAGAGAGATTCAGCAATGAGTGTCT 339
DB 5291 TATTAATTTACTTATTAATTAATTTATTTATTTATTTATTTATTTATTTATTT 5232
```

QY 340 TACAAGTGGAAAAACAAGTTAACTTAAGTGAACCCCTCCCTTGACAGATCAA 392
DB 5231 AACAACAAACAAAACAAAACAAAACAAACCCCAATTTCACACCTATATCTA 5179

RESULT 13

US-09-925-065A-121976
; Sequence 121976, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 121976
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-121976

Query Match
Best Local Similarity 8.0%; Score 40.8; DB 13; Length 546;
Matches 93; Conservative 1; Mismatches 88; Indels 0; Gaps 0;

QY 106 TTCTCCAGTTGTATTAAGCCCTCCACAGATTAAGTCCCTATGCCAACAAGAAAG 165
DB 104 TATGTAAAGTGTATGAATGCTGAGGTAGACATGAGTAAGAGAAAAAATATTAAT 163
QY 166 TCAATACATCTCTCTAGTCTCATTTATTTTCATTAGATAGCCGTTTTTACTACAAC 225
DB 164 TTAATAATTTCTCTGAGAAATGTTATTTATTAAGAAAAATGTAGTACATTTCTTA 223
QY 226 TCAATTAAGATGAACAGATGAATGGCTTAGTGCATGTTTATTAAGAGATTAATAA 285
DB 224 CATCATAGACAGTAATGAATGTACGATGTATGATGATTAATAACATTTAGAAATTAAGA 283
QY 286 TA 287
DB 284 GA 285

RESULT 14

US-10-172-086-13/C
; Sequence 13, Application US/10172086
; Publication No. US20030113750A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; FILE REFERENCE: of prostate tumors
; CURRENT APPLICATION NUMBER: US/10/172,086
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 13
; LENGTH: 7040
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-172-086-13

Query Match
Best Local Similarity 7.9%; Score 40.6; DB 16; Length 7040;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 101 TACGTTTTCTCAGTTGTATTAAGACCCCTCCACAGATTAAGTCTATGCAACAAAGA 160
DB 3586 TATATTTTAAAAAATCAATTAATTAATTCCTTTCTTAACACCTATATTTAAAT 3527
QY 161 AATGTCATATCATTTCTTAGTCTCATTTATTTTCATTAGATAGCCGTTTTTACT 220
DB 3526 ATATATAACCATTTTATATATTAACAATTCATTTTAAAAATATTAACCTTAAAAA 3467
QY 221 ACACTCAATTAAGATGAACAGATGAATGGGTTAGTACTGTTTATTAAGAAGTAT 280
DB 3466 ATAAATTAATAACATAACCAATTTACAAAAACAATTTAAATCTTAACCCCTAAAT 3407
QY 281 AAGATTAAT 291
DB 3406 AAAATTTCTAT 3396

RESULT 15

US-10-221-714A-161/C
; Sequence 161, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: tumor suppressor genes and oncogenes
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 161
; LENGTH: 7040
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-161

Query Match
Best Local Similarity 7.9%; Score 40.6; DB 19; Length 7040;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 101 TACGTTTTCTCAGTTGTATTAAGACCCCTCCACAGATTAAGTCTATGCAACAAAGA 160
DB 3586 TATATTTTAAAAAATCAATTAATTAATTCCTTTCTTAACACCTATATTTAAAT 3527
QY 161 AATGTCATATCATTTCTTAGTCTCATTTATTTTCATTAGATAGCCGTTTTTACT 220
DB 3526 ATATATAACCATTTTATATATTAACAATTCATTTTAAAAATATTAACCTTAAAAA 3467
QY 221 ACACTCAATTAAGATGAACAGATGAATGGGTTAGTACTGTTTATTAAGAAGTAT 280
DB 3466 ATAAATTAATAACATAACCAATTTACAAAAACAATTTAAATCTTAACCCCTAAAT 3407

Qy	281	AAAGATACCTAT	291
Db	3406	AAATTCTTAT	3396

Search completed: November 21, 2005, 11:54:58
Job time : 822:717 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 05:12:21 ; Search time 126.169 Seconds
(without alignments)
6639.056 Million cell updates/sec

Title: US-10-099-663-1_COPY_1115_1626

Perfect score: 512
Sequence: 1 cagtgatcctctcatctcatc.....gtctgctacagacagaag 512

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.4	8.9	1141	US-09-806-708B-22	Sequence 22, Appl
2	42.2	8.2	1257	US-09-270-767-15058	Sequence 15058, A
3	40	7.8	733	US-09-270-767-10136	Sequence 10136, A
4	39.2	7.7	4860	US-09-270-767-10504	Sequence 10504, A
5	38	7.4	640681	US-09-790-988-1	Sequence 1, Appl1
6	37.8	7.4	7218	US-08-232-463-14	Sequence 14, Appl1
7	37.6	7.3	601	US-09-949-016-125489	Sequence 125489, A
8	37.6	7.3	44353	US-09-949-016-15302	Sequence 15302, A
9	37.4	7.3	663	US-08-998-416-187	Sequence 187, App
10	37.4	7.3	696	US-08-998-416-779	Sequence 779, App
11	37.4	7.3	719	US-08-998-416-1138	Sequence 1138, App
12	37.4	7.3	856	US-08-998-416-289	Sequence 289, App
13	37	7.2	832	US-09-621-976-2813	Sequence 2813, App
14	36.8	7.2	43117	US-09-949-016-17589	Sequence 17589, A
15	36.6	7.1	601	US-09-949-016-16507	Sequence 16507, A
16	36.6	7.1	601	US-09-949-016-16564	Sequence 16564, A
17	36.6	7.1	27543	US-09-949-016-17455	Sequence 17455, A
18	36.6	7.1	36643	US-09-949-016-11860	Sequence 11860, A
19	36.6	7.1	36821	US-09-949-016-11860	Sequence 11860, A
20	36.6	7.1	36821	US-09-949-016-16404	Sequence 16404, A
21	36.2	7.1	601	US-09-949-016-173718	Sequence 173718, A
22	36.2	7.1	601	US-09-949-016-173765	Sequence 173765, A
23	36.2	7.1	832	US-09-621-976-2813	Sequence 2813, App
24	36.2	7.1	1587	US-09-710-279-1929	Sequence 1929, App
25	36.2	7.1	1878	US-09-134-001C-1667	Sequence 1667, App
26	36.2	7.1	3000	US-09-710-279-3907	Sequence 3907, App
27	36.2	7.1	3368	US-09-710-279-4169	Sequence 4169, App

28	36.2	7.1	81819	US-09-949-016-16661	Sequence 16661, A
29	36.2	7.1	81819	US-09-949-016-16662	Sequence 16662, A
30	36	7.0	1055	US-09-806-708B-23	Sequence 23, Appl
31	36	7.0	168394	US-09-949-016-13002	Sequence 13002, A
32	36	7.0	786431	US-09-751-389-3	Sequence 3, Appl1
33	35.8	7.0	1557	US-09-270-767-28519	Sequence 28519, A
34	35.8	7.0	3128	US-09-270-767-12706	Sequence 12706, A
35	35.8	7.0	192956	US-09-949-016-14382	Sequence 14382, A
36	35.8	7.0	1230025	US-09-438-452A-1	Sequence 1, Appl1
37	35.8	7.0	1230230	US-09-438-185A-1	Sequence 1, Appl1
38	35.6	7.0	601	US-09-949-016-44395	Sequence 44395, A
39	35.6	7.0	601	US-09-949-016-58898	Sequence 2845, App
40	35.6	7.0	1137	US-09-107-532A-2846	Sequence 2846, App
41	35.6	7.0	139562	US-09-949-016-13451	Sequence 13451, A
42	35.6	7.0	256171	US-09-949-016-12822	Sequence 12822, A
43	35.6	7.0	256176	US-09-949-016-15524	Sequence 15524, A
44	35.4	6.9	601	US-09-949-016-121946	Sequence 121946, A
45	35.4	6.9	601	US-09-949-016-202243	Sequence 202243, A

ALIGNMENTS

```
RESULT 1
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806, 708B
; PRIOR APPLICATION NUMBER: 2001-04-03
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. PAEI promoters
US-09-806-708B-22

Query Match      8.9%; Score 45.4; DB 4; Length 1141;
Best Local Similarity 9.9%; Pred.No. 0.003;
Matches 46; Conservative 197; Mismatches 221; Indels 0; Gaps 0;

QY 12 TCATCCATGCTTATATGAGTAGATGATTCCTCCAAATAGATGAGCAAGC 71
    :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 110 YBMTWNNKKGKQWRRYRMBDVTVDHYTANNNTWTCMMXDKDKRMMWKK 169
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 72 AGAGATTTGTTTATCTGTTGGTAAATAGCTTTTCCAGTTGTATTAAGACCTCC 131
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 170 NNATGDDDTKXMMNNNNNCBVTWVVRKYTRDWBKEMNYGMBWKKWSYDVTYVWV 229
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 132 CACCGATATAGTCTATGCAAGAAATGCAATACATTCCTAGTCTCTAT 191
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 230 WDMCRKRVRRVTRGRMRNTVAVBTYHRRRRYNGWBTBAAVYRWYNNNNNNKAKC 289
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 192 TATTTTCATTAGTAGCCGTTTCTTACTACATCAATTAAGATGAGCAATGATG 251
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 290 KKAQWGNRABVNSCTTWKSKTKTYRSCVANNRABADAKDKHKKWSAAMGYWNN 349
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 252 GTTAGTACTGTTTAAAGAGATTAATTAAGATATCTATCATTTGAGCAATGAG 311
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 350 NNNNNNTYKKBABRABMDVWHSAMKWMANAAHYRKKWTBYKRTVNNNNNGTTMMKR 409
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 312 GAGGAGATTCAGCAAAAGTGTGCTTCAAGTGAAGAAACAGTTAACTAACTGAC 371
    :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```

Dd      410 MMAMWYKMDMBGTYNNNNNGRYYGWTXXKXMMYYTKXKANNCKWAPMDHKCTHNN 469
OY      372 CCCCCTCCCTGACAGATCATGACCAAGTGGCTTTAGCCAGCACATCATCATGTA 431
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      470 TTMWKKTKTYNNNCYKMSMTNGSHBBAALVYTMWMMWRYYAHANNNDYMKKCTWYK 529
OY      432 ATTGCTTTCGATAGACCTGTTCATTAATTCCTTTGGCAAGC 475
           : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      530 YBVCSKMNNNYAAWTKSSMNTSYRYRWKTNNSMRWRSDTRSM 573

RESULT 2
US-09-270-767-15058/c
; Sequence 15058, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15058
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15058

```

	Query Match	Similarity	8.2%	Score 42.2	DB 4	Length 1257	
	Best Local	Similarity	53.3%	Pred. No. 0.026	Mismatches 78	Indels 0	Gaps 0
	Matches	89	Conservative	0			
Qy	73	GAGATTGGTTTATCTGTTGGTAAATACCTTTCTCAGTTGTATTAAGACCTGCC	132				
Db	364	GAGTTTATGATTATATTTTAAGTATTTCACGTCAACACGACACGACCTACTC	305				
Qy	133	ACCAATATTAAGTCTTATGCACAAAGAAATGTCAATCAATCTTGTGTCATTATT	192				
Db	304	ACAGCAAAAACGTACAGAGAGAAAAGAAAGATATAAAAAGATGTTCTTCACAAATA	245				
Qy	193	ATTTCATTAGTAGCCGCTTTTCTACTCAACTCAAAATTAAGTAA	239				
Db	244	TGTTTATAGCATTAAGGTGTGGCCATTCATATCAATATATAAGTA	198				

```

RESULT 3
US-09-270-767-10136/C
; Sequence 10136, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10136
; LENGTH: 733
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-10136

```

Query Match	7.8%	Score	40	DB	4	Length	733
Best Local Similarity	50.5%	Pident	No	0.092			
Matches	97	Conservative	0	Mismatches	95	Indels	0
				Gaps	0		
Oy	48	TCACCAATTAGAAATGCAACAAACAGAGATTGTGTTTATCTGTTGGTAAATACGTTT	107				
Db	306	TAAATAAATAGTCAAGTTTATATTGAGTTTATATGATTTATATTTTAAAGTTATTTTAACT	247				
Oy	108	TCTCCAGTTGTATTAAGACCTCCACACAGTTTAAAGTCTCTATGCAACAAAGAAATGTC	167				

Db	246	GCAACGCCACGACGACGACCTCTACACGCAAAAAACCTACAAAGAGAAAGAGGATA	187
Qy	168	AATACATCTCTTAGTCCTATTATTATTTCTATGATAGCCGTTTTTTACTCAATC	227
Db	186	AAAAAGTGATTTCTCTTCAATATGTTATGTGCAATAAAGGTGTGCCATTCATATC	127
Qy	228	AAATAGATGAA	239
Db	126	AAATATTAAGTA	115

```

RESULT 4
US-09-270-767-10504/C
; Sequence 10504, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10504
; LENGTH: 4860
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-10504

```

Query Match	7.7%	Score 39.2	DB 4	Length 4860
Best Local Similarity	46.6%	Pred. No. 0.33		
Matches	125	Conservative	0	Mismatches 143
			Indels	0
			Gaps	0
QY	97	TAAATAGCTTTTCCTCAGTGTGATPAAAGACCCCTCCACACAGTATPAAAGTCCCTATGCAACA	156	
Db	1390	TAAATCAGTATATATTTTGTGTAATAATATCAACAAGCATAAAATTAATGGACATATATAAT	1333	
QY	157	AAGAAATATGCAATACATTCCTTAGTCATTTATTTTCATTAGATAGCCGGTTT	216	
Db	1330	AATTAATTTATTTATTTTAAATATGCGCATATGCTTTAAATTTTTCGCTTTCA	1271	
QY	217	TACTACAACTCAAAATPAGATGACGAATGATGGGTTAGTACTGTTTATPAAAGAGAG	276	
Db	1270	TTTGTGCGCATPAAATTTGTTAAATTTTACATTTGATAGCTTTGTATTTTAAAGAAATGC	1211	
QY	277	TAATPAAAGATCTATCATCTTTGAGCAATPAGGGAGGAGATTCAGCAAAACAGTGT	336	
Db	1210	CAAAATAGTTTTCCTTCATATAGGGTAAGAAAGATATTAATTAACCAATTAGTTC	1155	
QY	337	GCTTACAAAGTGAAACAGATTAACTA	364	
Db	1150	ATTTTATGTTTAAATPAAAGTCAGTCGA	1123	

```

RESULT 5
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: MATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHISA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENE/DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

```

LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match
Best Local Similarity 46.6%; Pred. No. 4.9;
Matches 122; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

7.4%; Score 38; DB 4; Length 640681;
Query 24 TATTATGTGAGTAGAATGATGTCACCAATTTAGAGACAAAGCAGAGATTTGTGT 83
Db 189923 TATTATTTCTGAGAAATGATGATGATTTCTCAAAATTTGATGATCGGAGATATAT 189982
QY 84 TTTATCTGTTGGTAATAGTTCTTCTCCAGTTGTAAGACCCCTCCACGATATAA 143
Db 189983 TATAGCAAAATTTAGAAATTTCTCTCCGATATATTTAAATTAAGAAAAATTA 190042
QY 144 GTCCATGACAAAGAAATGTCATATCTCTAGTCTCATTTATTTTCATTAG 203
Db 190043 TTTTCTTTTAAACAAAAATATTAATAATTTTGGTGTATTTTATCAAAATTA 190102
QY 204 ATAGCCGTTTTTTTACTACAACTCAATTAAGATGACAGATGATGATGCTGT 263
Db 190103 TTCTCTTTCTAGAAACAAATGATGATTTATTAAGATGATGTTTAAATAAAT 190162
QY 264 TTTTAAAGAGAGTAAATGA 285
Db 190163 AAAAAATGAACAAACAAATA 190184

RESULT 6

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZpdt-Fls
US-08-232-463-14

Query Match
Best Local Similarity 13.5%; Pred. No. 0.97;
Matches 30; Conservative 103; Mismatches 90; Indels 0; Gaps 0;

7.4%; Score 37.8; DB 1; Length 7218;
Query 148 TATGCAACAAAGAAATGTCATATCTCTAGTCTCATTTATTTTCATTAGATG 207
Db 1494 TAGGATCACTGATTAATTTACTTATCTGCAAGTAGTTAAGATTAAGAAATTTGGTAC 1435
QY 208 CCGGTTTTTACTACAACTCAATTAAGATGACAGATGATGATGCTGTTAT 267
Db 1434 RRR 1375
QY 268 AAGAAAGAGTAAATTAAGATCTATCTCATTTTGGCAATTAAGGAGGAGATTCAGC 327
Db 1374 RRR 1315
QY 328 AAGAGTGTGCTTACAGATGAAACAGATTAAGTGA 370
Db 1314 RRR 1272

RESULT 7

US-09-949-016-125489/c
Sequence 125489, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 125489
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-125489

Query Match
Best Local Similarity 49.0%; Pred. No. 0.42;
Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

7.3%; Score 37.6; DB 4; Length 601;
Query 148 TATGCAACAAAGAAATGTCATATCTCTAGTCTCATTTATTTTCATTAGATG 207
Db 365 TATTAATAATGATTAATATATACACAGATGATATTAACATGATTAAGTTAACTACATAT 306
QY 208 CCGGTTTTTACTACAACTCAATTAAGATGACAGATGATGCTGTTAT 267
Db 305 ACATRTATATAGTTAGTATATCTAGAAAGCATGCAAGATTTTCTGCTACTATCTAA 246
QY 268 AAGAAAGAGTAAATTAAGATCTATCTATTTGAGCATTAAGGAGGAGATTCAGC 327
Db 245 TTAATGCAATCTTATATAGTCTCTCAAAATTAATTAAGATTAATTAAGCGTCAAG 186
QY 328 AAGAGTGTGCTTACAGATGAA 351
Db 185 CAACATTATGCTTAAAGATTA 162

RESULT 8
US-09-949-016-15302/C
; Sequence 15302, Application US/09949016
; Patent No. 6812338
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15302
; LENGTH: 44353
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(44353)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15302

Query Match 7.3%; Score 37.6; DB 4; Length 44353;
Best Local Similarity 49.0%; Pred. No. 2.3;
Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 148 TATGCAACAAAGAAATGCAATACATCTCTAGTCATTTATTTTCATTAAATAG 207
DB 25736 TATTAATATAGTAATATATCAACAGATGATATTAACATGATATAGTTAACTACATAT 25677
QY 208 CCGGTTTTTACTACAACTCAATTAAGATGACAGATGATGGTTAGTGTAT 267
DB 25676 ACATGATATAGTAGTATATACAGAAAGATTTCTGGCTACTATCTAA 25617
QY 268 AAAGAAGAGTAATAAATATCTATCATCTTTGAGCAATAGGAGAGAGATTCACC 327
DB 25616 TTATGCACATCTTATAGCTCTCAAAATATTAAGATTAATATCAAGCGTCAGT 25557
QY 328 AAACAGTGCCTTACAAGTGGA 351
DB 25556 CAACATTATGCTTAAAAAGTAAAA 25533

RESULT 9
US-08-998-416-187/C
; Sequence 187, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPYII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 187:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 663 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1074UP
US-08-998-416-187

Query Match 7.3%; Score 37.4; DB 3; Length 663;
Best Local Similarity 58.6%; Pred. No. 0.5;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 180 TAGTCTATATATTTTTCATATAGATGCGGTTTTTACTCACTCAATTAAGATGA 239
DB 574 TAGATTTATATTAATCTTATTAAGTTAAACATTATTAATGATCAATTAATTAATA 515
QY 240 CAGATGATGAGGTAGTACTGTTATTAAGAGATTAAGATTAAGATTAAGATTA 290
DB 514 GGAATTAATTAAT 464

RESULT 10
US-08-998-416-779
; Sequence 779, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPYII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 779:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1501RP
US-08-998-416-779

Query Match
Best Local Similarity 58.6%; Score 37.4; DB 3; Length 696;
Pred. No. 0.51;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 180 TAGCTCATTATTATTTTCATTAGATAGCCGGTTTTTACTCACTCAATAGATGAA 239
DB 453 TAGATTATTATTAACTTATTAGTTAAACATTATTATTATGATATATATATATAA 512

QY 240 CAGATGAATGGGTTAGTACTGTTTATTAAGAAGATATAAGTACTA 290
DB 513 GGAATTAACATTATGATATATAGTTATTATAGAACCAATGAAATACTA 563

RESULT 11
US-08-998-416-1138/c
Sequence 1138, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtel, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1138:
SEQUENCE CHARACTERISTICS:
LENGTH: 719 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1692UP
US-08-998-416-1138

Query Match
Best Local Similarity 58.6%; Score 37.4; DB 3; Length 719;
Pred. No. 0.51;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 180 TAGCTCATTATTATTTTCATTAGATAGCCGGTTTTTACTCACTCAATAGATGAA 239
DB 574 TAGATTATTATTAACTTATTAGTTAAACATTATTATTATGATCATATAATATATAA 515

QY 240 CAGATGAATGGGTTAGTACTGTTTATTAAGAAGATATAAGTACTA 290
DB 514 GGAATTAACATTATGATATATAGTTATTATAGAACCAATGAAATACTA 464

RESULT 12
US-08-998-416-289/c
Sequence 289, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtel, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSEYII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 289:
SEQUENCE CHARACTERISTICS:
LENGTH: 856 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

Db	368	GAATGCTTCAATAGAAAATCAATCATGTATATC-ACCACTATTAGAAAAGAA	426
Qy	314	GGAGAGATTACGAAACAGTGTCTTACAAGTGAAAAACAAGTTAACTTAAAGTACCC	373
Db	427	AAATCTAACAGATAGACCATCTGAATGAGAGCTGAGAAAGCATTAAGACAAATTCATTA	486
Qy	374	CCCTCCTTGACAAGATCAATATGCCACAGTTGA	404
Db	487	CTTTCTGTATTAATCATCTCAAGATATGA	517

Search completed: November 21, 2005, 08:29:54
 Job time : 131.189 secs

Blank (uspto)

may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). "

Query Match	9.3%	Score 47.6;	DB 9;	Length 1206;
Best Local Similarity	45.5%;	Pred. No. 0.16;		
Matches 170; Conservative	0;	Mismatches 204;	Indels 0;	Gaps 0

Oy	11	TTCAATCTAAGTGTATTATATGAGAGTAAGAATAGATTCTCCAAATTGAATGCACAAG	70
Dd	999	TTTAACCTATATATTTTTTAAATGAATACTATCTAATTCATAAAGAAATTAACCTATAG	940
Oy	71	CAGAGATTGGTGTATTATCTGTGGGTAATACGTTTTCTCAGTGATATAAGACCCTC	130
Dd	939	CTTATATATTAATATATATTTAAATNTCATATTTAATTAATTAAGATRTAAACTGTATA	880
Oy	131	CCACCAATATAAAGTCTATGCAACAAAAGAAATGTCAATACATTCTCTAGTCTCATTA	190
Dd	879	TAAATATATTAAATTAATTCATATAAAGTATATTATATAGATTAAATTAAGTACATTA	820
Oy	191	TTATTTTCATTGATAGGCCGGTTTTTACTACAACTCAAATTAAGTAGAACAGATGAATG	250
Dd	819	ATTATAAATATATTTAATATATATTTTAAATTAATFACATAAAAAAACTCAATTAATGTTAA	760
Oy	251	GGTATGTAAGTGTATTATAAGAAGATATAAAGATACATCATCATTTGAGACAAATAG	310
Dd	759	TATATGAATTTCCATPAACAAATATACRTAATTAATATAGATRAATTAATGCAAAATA	700
Oy	311	GGAGGAGAGATTTCAGCAAAACAGTGTCTTCAAGTGAAGAAAACAAGTTAACTAAAGTGA	370
Dd	699	CAAAATTAATATATAAAAAATTATATATCAAAATTAATTAATTTACATPAACAAATATA	640
Oy	371	CCCCCTCCTTGAC	384
Dd	639	CACACTTCATTTAC	626

RESULT 2	
CNS014PQ	
LOCUS	
DEFINITION	CNS014PQ 987 bp DNA linear GSS 26-JUL-1999 Drosophila melanogaster genome survey sequence SP6 end of BAC BACN12922 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL104456
VERSION	ALI04456.1 GI:5616067
SOURCE	GSS.
ORGANISM	Drosophila melanogaster (fruit fly) Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyndroidea; Drosophilidae; Drosophila. 1 (bases 1 to 987) Genoscope.
REFERENCE	Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
AUTHORS	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CBPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pbeloBAC11.
TITLE	
JOURNAL	
COMMENT	

```

FEATURES
source
location/Qualifiers
1..987
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN12P22"
/clone_1fb="DIOSBAC"
/plasmid="pbeloBAC11"

```

```
ORIGIN
/note="end : SP6"
```

Query Match	9.0%	Score 46	DB 9	Length 987
Best Local Similarity	33.5%	Pred. No. 0.4		
Matches 78	Conservative 51	Mismatches 104	Indels 0	Gaps 0

[illegible]

	RESULT 3	
CNS00E0E/c		
LOCUS		
DEFINITION		
CNS00E0E	1101 bp	DNA linear GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #		
BACR2A22 of RPl1-98 library from Drosophila melanogaster (fruit		
fly), genomic survey sequence.		

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AL067626	AL067626.1	GI:4948726	GSS.	<i>Drosophila melanogaster</i> (fruit fly)	
				<i>Drosophila melanogaster</i>	
				Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota	
				Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
				Phyldiroidea; Drosophilidae; Drosophila.	
				1 (bases 1 to 1101)	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1101)		Genoscope.		
		Direct Submission		
		Submitted (02-JUN-1999)	Genoscope - Centre National de Sequencage	
		Bp 191 91006 EVRY cedex - FRANCE (E-mail : seguet@genoscope.cns.fr)		
		- Web : www.genoscope.cns.fr		
		Determination of this BAC-end sequence was carried out as part of a		

COMMENT

- Web : www.genoscope.cns.it/
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oosawa and Aaron Mammaster in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

FEATURES
    source
        location/Qualifiers
            1..1101
            /organism="Hirosonchila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="Enson:7227"
            /clone="BACR28A2"
            /clone_id="RPC1-98"
            /note="end : TET3"
ORIGIN

```

Query Match	8.8%	Score	45.2	DB	9	Length	1101
Best Local Similarity	41.8%	Pred. No.	0.67				
Matches	79	Conservative	28	Mismatches	82	Indels	0
						Gaps	0

MON NOV 21 12:27:20 2005

RESULT 4	
AZ541558/c	
LOCUS	AZ541558 910 bp DNA linear GSS 14-NOV-2000
DEFINITION	ENTEB15TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION	AZ541558
VERSION	AZ541558.1
KEYWORDS	GI:11149198
SOURCE	GSS.
ORGANISM	Entamoeba histolytica
	Entamoeba histolytica
	Eukaryota; Entamoebidae; Entamoeba.
REFERENCE	1 (bases 1 to 910)
AUTHORS	Loftus,B., Van Aken,S. and Fraser,C.
TITLE	Determination of clone end sequences from Entamoeba histolytica HM:IMSS sheared DNA library
JOURNAL	Unpublished (2000)
COMMENT	Contact: Brendan J Loftus

JOURNAL
 COMMENT
 Unpublished (2000)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjoftus@tigr.org
 Clones are derived from the *Entamoeba histolytica* HMI.1MS5 sheared
 DNA library
 Seq primer: M13-Forward
 Class: Shotgun
 High quality sequence start: 26
 High quality sequence stop: 840.
 Location/Qualifiers
 1..910
 FEATURES
 SOURCE

```

/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHDOS1, Site 1: Bst I, Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1993)."

```

ORIGIN

Query Match	8.7%	Score 44.4;	DB 8;	Length 910;
Best Local Similarity	46.7%;	Pred. No. 1;		
Matches 141;	Conservative 0;	Mismatches 161;	Indels 0;	Gaps 0;
37	AGATTAGATTTCACCAATTAGATGCAACGACGAGATTTCGTTTATCTGTGGG	96		

Db 310 AGAATTAATAAGAAATAGATTAATAATGAATAAACAATTATTTGGTATATTTCATGAAAT 250

Qy 97 TAAATAGCTTTTCTCCAGTTGTATATAAGACCCTCCACAGATATAAAGTCTTATGCACA 156

Db 250 TAATTCATTAATGAATAAATAGATTAAGATTAATAGAAATTAAGATTAACAATTAAGAAAT 190

Qy 157 AAGAAATAGTCAATACATCTCTTATGCTCATTTATTTATTTATTAAGATAGCCGTTT 216

Db 190 GATAAAATACATCAACATATTTATGTATTATTTCTTTATCTTTATTCATCTTATTTTC 131

Qy 217 TACTTCAACTCAAAATTAAGATGAACGAATGAATGGTTAGTACTGTTTATATAAGAAAG 276

Db 130 TATTTAACAATTAATTAACAACAACATTAATMAAAGATTAATAGATTAACACATCATTA 71

Qy 277 TAATTAAGATCTATCATCATTTTGAGGCAMTAAGGAGGAGAGATTCAGCAAAACGTGT 336

Db 70 AATAAAGATTAAGATTAATAATATATACAACAATATATAAACAAGAAATACCAAAATACTTT 11

Qy 337 GC 338

Db 10 CC 9

```

RESULT 5
BH155022
LOCUS
      BH155022
DEFINITION
      BH155022
      ENTS155Tf Entamoeba histolytica Sheared DNA Entamoeba histolytica
      genomic, genomic survey sequence.
      894 bp      DNA      linear      GSS 24-SEP-200
ACCESSION
      BH155022
VERSION
      BH155022.1
KEYWORDS
      GI:15726739
SOURCE
      GSS.
      Entamoeba histolytica
      Entamoeba histolytica
      Entamoeba histolytica
      Eukaryota; Entamoebidae; Entamoeba.
      1 (bases 1 to 894)
REFERENCE
      Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
      Determination of clone end sequences from Entamoeba histolytica
      HMI:IMSS sheared DNA library (2001)
      Unpublished (2001)
JOURNAL
      Comment: Brendan J Loftus
      Department of Eukaryotic Genomics
      The Institute for Genomic Research
      9712 Medical Center Dr., Rockville, MD 20850, USA
      Tel: 301 838 0208
      Fax: 301 838 3543
      Email: b1ofus@igf.org
      Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
      DNA library
      Seq primer: M13-Forward
      Class: shotgun
      High quality sequence start: 52
      High quality sequence stop: 822.
      Location/Qualifiers
          1..894
FEATURES
      source

```

```

/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:1MS3"
/db_xref="taxon:5759"
/cClone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pMOSt1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broch cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. in Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Bartell, Oxford University Press, 1999)."

```


Db 918 DDCAGDXXDDGKXKADDDTDTGTXDDDDKXKMDMDKXAGTGDATAMANAATDMMWGMWA 859

QY 298 TTGAGGCATTAAGGAGGAGAGATTGACCAACAGTGTGCTTACAGTGGAACAAACAGT 357

Db 858 DADWMTMDAADMDWADDDMDWADMDWAMWAGARTADRDWDGDRAGKGRKGRDRBK 799

QY 358 TAAACTTAAGTGA 370

Db 798 RADDKRDADDDRD 786

RESULT 8
CL108838/c
LOCUS 1158 bp DNA linear GSS 05-JAN-2004

DEFINITION ISB1-5012, Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-5012,

ACCESSION CL108838

VERSION CL108838

KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

REFERENCE 1 (bases 1 to 1158)

AUTHORS Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

TITLE A physical map of the xenopus tropicalis genome

JOURNAL Unpublished (2003)

COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: subraisons@wustl.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 36
High quality sequence stop: 766.

FEATURES
source
1. 1158
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-5012"
/note="Vector: pBelobAC11, ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN
Query Match 8.6%; Score 43.8; DB 9; Length 1158;
Best Local Similarity 52.5%; Pred. No. 1.6;
Matches 96; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 115 TTGTAAAGACCCCTCCACGATATAAGTCTCTATGACCAAGAAATGTCAATACAT 174

Db 781 TGGAAATAGGCATCTCAGCAATAGTAGGCTTAAGAAAAAGAAAAAGAAAAA 722

QY 175 TCTCTAGTCTCATTTATTTTTCATTAAGATAGCCGGTTTTTTTACTACACTCAATTAAG 234

Db 721 GCGTTGGGTATTTGTTAATGGTTTATAGAGCCGTTAAGCCACAGGTTAAAGAAA 662

QY 235 ATGACAGATGATGATGGTTAGTACTGTTTATTAAGAAAGATTAATTAAGATCTATCAT 294

Db 661 GTTAAATATCTAATATGTTATGACATATCCCATAGTCAATATCTTATTAATACCT 602

QY 295 CAT 297

Db 601 TAT 599

RESULT 9
CG769354
LOCUS 927 bp DNA linear GSS 29-OCT-2003

DEFINITION TcB4.1_H04_Sp6 Tribolium BAC library Tribolium castaneum genomic,
genomic survey sequence.

ACCESSION CG769354

VERSION CG769354.1 GI:38022533

KEYWORDS GSS.

SOURCE Tribolium castaneum (red flour beetle)

ORGANISM Tribolium castaneum

REFERENCE 1 (bases 1 to 927)

AUTHORS Savard, J. and Tautz, D.

TITLE Tribolium castaneum BAC-ends sequencing project

JOURNAL Unpublished (2003)

COMMENT Contact: Savard, J.
Abteilung fur Evolutionsgenetik, AG Tautz
Institut fur Genetik, Universitat zu Koln
Weyertal 121, 50931 Koln, Germany
Tel: 49 221 470 6911
Fax: 49 221 470 5975
Email: savard@uni-koeln.de
Class: BAC ends.

FEATURES
source
1. 927
Location/Qualifiers
/organism="Tribolium castaneum"
/mol_type="genomic DNA"
/strain="GA-2"
/db_xref="taxon:7070"
/clone="Tribolium BAC library"
/note="Vector: pBAC3.6; Site 1: EcoRI, Site 2: EcoRI;
Library constructed by Exelixis Inc."

ORIGIN
Query Match 8.5%; Score 43.6; DB 9; Length 927;
Best Local Similarity 53.8%; Pred. No. 1.7;
Matches 112; Conservative 0; Mismatches 94; Indels 2; Gaps 1;

QY 138 TTATAAGTCTTATGCACAAAGAAATGTCAATACCTCTTCTAGTCTCATTTATTTT 197

Db 26 TATTAAGTCAGTGTAAATTTAAACTGTCAATTAATTTATTTCTTACAAATGTCAACTA 85

QY 198 CATTAATAGCCGGTTTTTTTACTACACTCAATTAAGATGAGCAATGAATGGTAGT 257

Db 86 AATTCATTTTACACCTGACGACAAAGAAATCTTTATTTTAAACAAAGAACCGATATT 145

QY 258 GACTGTTTATTAAGAAAGATTAATAA--GATCTATCATCATTTGAGCAATTAAGGAGG 315

Db 146 AATCGCATTAATTAAGATTAATTAATTAACCGATACATGTTATTAAGATCAAGCTGTAG 205

QY 316 GAGAGATTGACGAACAGTGTCTTACA 343

Db 206 AAGAGATTACAGAAACCGCTGTTTCA 233

RESULT 10
AV008603 443 bp mRNA linear EST 25-AUG-1999

LOCUS AV008603

DEFINITION AV008603 Mus musculus 18-day embryo C57BL/6J Mus musculus cDNA

ACCESSION AV008603

VERSION AV008603.1 GI:4785590

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 443)

AUTHORS Carninci, P., Shibata, K., Oza, Y., Komu, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
Kawai, D., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H.,
Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y.,
Sugahara, Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y.,

tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999). "

ORIGIN

Query Match 8.4%; Score 43; DB 8; Length 879;
Best Local Similarity 49.8%; Pred. No. 2.4;
Matches 109; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 148 TATGCACAAAGAAATGTCATATCCTCTAGTCATTTATTTTCATTAGATG 207
DB 303 TACAAAAGAGTAATTCCTGGAATTAATGTTTATGAATGAATTCAG 244
QY 208 CCGGTTTTTACTACAACTCAATTAAGATGACAGATGATGGTTAGTCTTTAT 267
DB 243 AAGTGAATATCCACAGATGTAAGTAATAGAAATGATGTTTTCATATTCAT 184
QY 268 AAGAAAGATTAATTAAGATATCATCATTTGAGCAATTAAGAGAGAGATTGACG 327
DB 183 CATTAACAAATTAATATATCATCATTCATCATTTGAGTTGATGTTTATC 124
QY 328 AAACAGTGTCTTACAGATGGAACAAAGTTAACTAAA 366
DB 123 ATTGTGATGTGAAGAAATTAAGAAAGATTAACAA 85

RESULT 13

BH147340/C BH147340 943 bp DNA linear GSS 27-AUG-2001
LOCUS ENTOM341R Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, genomic survey sequence.

ACCESSION BH147340
VERSION BH147340.1 GI:15304295
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 943)
AUTHORS Loftus B., Wang Z., Van Aken S. and Fraser C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seg primer: M13-Reverse
Class: Shotgun
High quality sequence start: 60
High quality sequence stop: 602.
Location/Qualifiers

FEATURES

1..943
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1, Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999). "

ORIGIN

Query Match 8.4%; Score 43; DB 8; Length 943;
Best Local Similarity 49.8%; Pred. No. 2.4;
Matches 109; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 148 TATGCACAAAGAAATGTCATATCCTCTAGTCATTTATTTTCATTAGATG 207
DB 575 TACAAAAGAGTAATTCCTGGAATTAATGTTTATGAATGAATTCAG 516
QY 208 CCGGTTTTTACTACAACTCAATTAAGATGACAGATGATGGTTAGTCTTTAT 267
DB 515 AAGTGAATATCCACAGATGTAAGTAATAGAAATGATGTTTTCATATTCAT 456
QY 268 AAGAAAGATTAATTAAGATATCATCATTTGAGCAATTAAGAGAGAGATTGACG 327
DB 455 CATTAACAAATTAATATATCATCATTCATCATTTGAGTTGATGTTTATC 396
QY 328 AAACAGTGTCTTACAGATGGAACAAAGTTAACTAAA 366
DB 395 ATTGTGATGTGAAGAAATTAAGAAAGATTAACAA 357

RESULT 14

CNS0026Z CNS0026Z 1101 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN01A10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL097301 GI:5608912
VERSION AL097301.1
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-Jul-1999) Genoscope - Centre National de Sequencage:
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web: www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.biol.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES

1..1101
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN01A10"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : 77"

ORIGIN

Query Match 8.3%; Score 42.6; DB 9; Length 1101;
Best Local Similarity 29.0%; Pred. No. 3.2;
Matches 108; Conservative 92; Mismatches 170; Indels 3; Gaps 1;
QY 37 AGAATAGATTCACCAATTAATGAATGCAAGAGAGATTGGTTTATCTGTGGG 96

Db	620	AAAAAABRTATMTTATATATATATAATAAATAAAAAAATATTCTTCKTMAAKTTMAAK	679
Qy	97	TAATAGCTTTCTCCAGTTGTATTAAGACCTCCCAAGATATAAGTCTTATGCACA	156
Db	680	MAMAMMTTKTTTTMMRAK--TKAAACAHAAPAKAKABAAAAAATWTAKABMA	736
Qy	157	AAGAAATGTCAATACCTCTGAGTCATTAATTTTCAATAGATAGCGGTTTT	216
Db	737	WMAMTWTAKTTHDTATTAMCTMTMAAATATTTTTTTTMMAMCKTMAAABTTTTTW	796
Qy	217	TACTACAATCAATTAAGATGAACAAGATGAATGGTTAGTACTGTTATTAAGAAG	276
Db	797	TATTTATTTTTTBEKYAAMACTCMAAAAAAAMAKKYDTAAKTYDAMMMMAAKATYA	856
Qy	277	TAATTAAGATCTATCATTCATTTTGAGGCAATAGGAGGAGAGATTGACAAACAGTG	336
Db	857	WTMTMTWAMWMAAATATWTATTAACKMSYCBAYACNMDDAMMMIMCCBACACCCNC	916
Qy	337	GCTTCAAGTGGAAAAACAAGTTAAACTAAAGTGAACCCCTCTTGAACAAGTCAATGCC	396
Db	917	YKNYAMMMHADKAAAAACHDKHMWMBAAAKKKBKBYMTMMWMTTATWTWTYCTWMA	976
Qy	397	ACAGTTGAGCTTT	409
Db	977	MMADTAMACHTW	989

RESULT 15			
CR684869/c			
LOCUS	CR684869	1201 bp	mRNA
DEFINITION	Tetradon nigroviridis full-length cDNA.		
			linear
			HTC 19-AUG-2004

ACCESSION	CR684863	GI:51182776
VERSION	CR684863.1	
KEYWORDS	HTC, cDNA, full-length, Tetraodon nigroviridis	
SOURCE	Tetraodon nigroviridis	
ORGANISM	Tetraodon nigroviridis	

REFERENCE 1 (bases 1 to 1201)

Genoscope - Centre National de Sequencage -
Submitted (10-AUG-2004)

COMMENT The sequences are based on single pass reads.

<http://www.genoscope.cns.fr/tetraodon>.

source 1. .1201

ORIGIN

Query Match	8.3%	Score 42.6;	DB 3;	Length 1201;
Best Local Similarity	49.8%	Pred. No. 3.3;		
Matches 108; Conservative	0;	Mismatches 109;	Indels 0;	Gaps 0;

Qy	153	TACCAAGAAATGTCATTCATCTCTAGTCGTCATTATTTTCATTAGATAGCCGG	212
Db	344	AAAAAAAAAAAAAAAAATATAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT	285
Qy	213	TTTTTACTCACTCAATTAAGATGAAGAAATGATGGGTAGTGACTGTTATATAAG	272
Db	284	AAAAAAAAAAAAAGTAGTTGGGTAAATTAATAATTAAAAAGTTGTTTTTAAAAAG	225
Qy	273	AGAGTATATTAAGTATCATCATTTGAGGCAATAGGAGGAGAGATTCAGCAAA	332
Db	224	AAAAAAAAAAGAAAAAAGTAAATTTAAAAAAGGGGAAAGGAGGAGGAAAAAG	165

Qy 333 GGTGCTTACAAGTGGAATAACCTAAAGTC 365
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 TTGTTTAAAAATTAAAAGGAGAAGAAAAAGGGG 128

Search completed: November 21, 2005, 08:26:16
Job time : 3105.36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 01:59:39 ; Search time 405.736 Seconds
(without alignments)
7470.150 Million cell updates/sec

```

Title:      US-10-099-663-1_COPY_1115_1626
Perfect score: 512
Sequence:   1 cagtgatcctcatctcat.....gtctgctacagacagaag 512

```

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

```

13: 1: geneseqn2004bs: *
12: 2: geneseqn2004as: *
11: 3: geneseqn2003cs: *
10: 4: geneseqn2003bs: *
9: 5: geneseqn2003as: *
8: 6: geneseqn2002bs: *
7: 7: geneseqn2002as: *
6: 8: geneseqn2001bs: *
5: 9: geneseqn2001as: *
4: 10: geneseqn2000bs: *
3: 11: geneseqn2000as: *
2: 12: geneseqn1998as: *
1: 13: geneseqn1890as: *
N_genseq_18dec04: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	512	100.0	2381	11	ADL90127	ADL90127 Chicken f
2	326	63.7	236	11	ADL90128	ADL90128 Chicken f
3	42.6	8.3	2000	8	ADA71538	Ada71538 Rice gene
4	42.4	8.3	5908	8	AAS451587	AAS451587 Chemical1
5	42.4	8.3	5908	6	ABK28232	ABK28232 DNA trans
6	41.8	8.2	10048	6	ABL70313	ABL70313 Chemical1
7	41.8	8.2	10048	6	AAS61251	AAS61251 Human
8	40.6	7.9	7040	4	AAS46439	AAS46439 Tumour su
9	40.6	7.9	7040	6	ABK33963	ABK33963 Human DNA
10	40.6	7.9	7040	8	ABZ10179	ABZ10179 Haematopo
11	40.6	7.9	7040	8	ABZ10033	ABZ10033 Haematopo
12	40.6	7.9	7040	8	ADA20348	ADA20348 Prostate
13	40.6	7.9	7040	8	ADA84155	ADA84155 Human ren
14	40.6	7.9	7040	10	ADBE84191	ADBE84191 Human lym
15	40.6	7.9	7040	10	ADBE84191	ADBE84191 Human lym
16	40.6	7.9	14095	6	ABL32477	ABL32477 Human imm
17	40.2	7.9	6065	6	ABK31356	ABK31356 Signal tr
18	40.2	7.9	6065	6	ABL70579	ABL70579 Chemical1
19	40.2	7.9	8365	6	AAS61260	AAS61260 Human gen
20	40.2	7.9	8365	2	AAS20056	AAS20056 Plasmodiu

C	21	39.4	7.7	5656	6	ABR28419	Abi28419 DNA trans
C	22	39.4	7.7	110000	10	ACF67367_23	Continuation (24
C	23	39.4	7.7	110000	10	ACF65386_5	Continuation (6
C	24	39.2	7.7	9117	6	ABLJ3069	AbiJ3069 Human imm
C	25	39	7.6	2933	4	ABLJ18616	Abi18616 Drosophi1
C	26	39	7.6	5311	6	ABLJ33019	AbiJ33019 Human imm
C	27	39	7.6	6240	10	ADD49062	Add49062 Human NO
C	28	39	7.6	22052	4	ABL20542	Abi20542 Drosophi1
C	29	39	7.6	26097	4	ABL03696	Abi03696 Drosophi1
C	30	39	7.6	36194	4	ABL20582	Abi20582 Drosophi1
C	31	39	7.6	59967	4	ABL15492	Abi15492 Drosophi1
C	32	38.4	7.5	32392	6	ABL56203	Abi56203 AMEPV gen
C	33	38.4	7.5	50000	6	ABL55643	Abi55643 AMEPV gen
C	34	38.2	7.5	60722	6	ABLJ32031	AbiJ32031 Human imm
C	35	38.2	7.5	56737	6	ABS69895	AbS69895 Human hyp
C	36	38.2	7.5	110000	12	ADH69807_4	Continuation (5
C	37	38	7.4	6802	6	ABLJ3231	AbiJ3231 Human imm
C	38	38	7.4	110000	6	ABA92787_1	Continuation (2 of
C	39	37.8	7.4	3322	4	ABL10846	Abi10846 Drosophi1
C	40	37.8	7.4	17527	6	ABLJ33433	AbiJ33433 Human imm
C	41	37.8	7.4	17527	6	AAS63333	Aas63333 Chemical1
C	42	37.6	7.3	3296	4	ABL17264	Abi17264 Drosophi1
C	43	37.6	7.3	8648	6	ABKJ1354	AbKJ1354 Signal tr
C	44	37.6	7.3	8648	6	ABL70573	Abi70573 Chemical1
C	45	37.6	7.3	13573	6	ABLJ33668	AbiJ33668 Human imm
ALIGNMENTS							
RESULT 1							
ID	ADL90127	standard; DNA; 2381 BP.					
XX	ADL90127;						
DT	20-MAY-2004	(first entry)					
DE	Chicken intestinal fatty acid binding protein, iFABP, gene, 5' region.						
XX	Chicken; de; intestinal fatty acid binding protein, iFABP;						
KM	gut specific promoter; transgenic.						
OS	Gallus gallus.						
XX	US2003177516-A1.						
PN	18-SEP-2003.						
PD	14-MAR-2002; 2002US-00099663.						
XX	14-MAR-2002; 2002US-00099663.						
PR	14-MAR-2002; 2002US-00099663.						
PA	(HORS/) HORSEMAN N D.						
XX	(PRA/) PRAAT S L.						
PI	Horseman ND, Pract SL;						
XX							
DR	WPI; 2003-898653/82.						
PT	New nucleic acid molecule comprising an isolated avian gut-specific gene						
PT	expression control region, useful for regulating heterologous nucleic						
PT	acids in transgenic avians, and for generating transgenic birds.						
XX	Claim 1; SEQ ID NO 1; 28bp; English.						
XX	The invention relates to an isolated nucleic acid comprising an isolated						
CC	avian gut-specific gene expression control region appearing as						
CC	ADL90127(Chicken intestinal fatty acid binding protein, iFABP, gene, 5'						
CC	region or ADL90128 (chicken iFAP promoter) or its degenerate variant.						
CC	Also included are a recombinant DNA molecule comprising an isolate avian						
CC	gut-specific gene expression control region operably linked to a nucleic						
CC	acid insert encoding a polypeptide, an expression vector that integrates						

CC into a host cell (and comprising the isolated avian gut-specific gene
CC expression control region), expressing a heterologous polypeptide in a
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
CC molecule, and culturing the transfected cell in a medium suitable for
CC expression of a heterologous polypeptide under the control of an avian
CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression
CC control region encoded by the recombinant DNA molecule), a eukaryotic
CC cell transformed with the expression vector for its progeny, which
CC expresses a heterologous polypeptide and a transgenic avian having a
CC heterologous polynucleotide sequence comprising the nucleic acid insert.
CC The nucleic acids are useful for regulating heterologous nucleic acids in
CC transgenic avians, as probes in nucleic acid hybridisation assays for
CC detecting the IFABP gene expression control region, and for generating
CC transgenic birds. The present sequence is the Chicken intestinal fatty
CC acid binding protein, IFABP, gene, 5' region.

XX SQ Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match 100.0%; Score 512; DB 11; Length 2381;
Best Local Similarity 100.0%; Pred. No. 2.8e-129;

Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGATCTCTCATCTCATCTGTTTATTTATGTTGAGATAGATTTCTGACCAATTAGA 60
DB 1115 CAGTGATCTCTCATCTCATCTGTTTATTTATGTTGAGATAGATTTCTGACCAATTAGA 1174
QY 61 ATGGACAAAGCAGAGATTTGTTTATCTGTTGGTAAATAGTTTCTCCAGTTGAT 120
DB 1175 ATGGACAAAGCAGAGATTTGTTTATCTGTTGGTAAATAGTTTCTCCAGTTGAT 1234
QY 121 AAAGACCTCCACCAGTAAAGTCTTATGCAACAAGAAATGTCATATCTCTT 180
DB 1235 AAAGACCTCCACCAGTAAAGTCTTATGCAACAAGAAATGTCATATCTCTT 1294
QY 181 AGTCTCATTTATTTTTCATTAGATAGCCGGTTTTTACTACACTCAATAATAGATGAAC 240
DB 1295 AGTCTCATTTATTTTTCATTAGATAGCCGGTTTTTACTACACTCAATAATAGATGAAC 1354
QY 241 AGAATAAGGGGTAGTGAAGCTGTTTAAAGAGATTAATGAATCATCTCATTTTG 300
DB 1355 AGAATAAGGGGTAGTGAAGCTGTTTAAAGAGATTAATGAATCATCTCATTTTG 1414
QY 301 AGGCAATTAAGGAGAGAGATTCAGCAACAAGTGTCTTACAAGTGAAGAAACAAGTTAA 360
DB 1415 AGGCAATTAAGGAGAGAGATTCAGCAACAAGTGTCTTACAAGTGAAGAAACAAGTTAA 1474
QY 361 ACTAAAGTGACCCCTCTCTTGAACAAGATCAATGCCACAGTTGAGCTTTAGCCAGCACA 420
DB 1475 ACTAAAGTGACCCCTCTCTTGAACAAGATCAATGCCACAGTTGAGCTTTAGCCAGCACA 1534
QY 421 TCATCATGTAATTTGCTTCCGTGATTAAGCCTGTTCAATTAATTCCTTTGCAAGCTCTGC 480
DB 1535 TCATCATGTAATTTGCTTCCGTGATTAAGCCTGTTCAATTAATTCCTTTGCAAGCTCTGC 1594
QY 481 TACTTACAGAGAGTGTGCTTACAGACAGAAAG 512
DB 1595 TACTTACAGAGAGTGTGCTTACAGACAGAAAG 1626

RESULT 2
ADL90128
ID ADL90128 standard; DNA; 336 BP.

XX ADL90128;

XX 20-MAY-2004 (first entry)

DB Chicken intestinal fatty acid binding protein, IFABP, gene, promoter.

XX Chicken, ds; intestinal fatty acid binding protein; IFABP;

KW gut specific promoter; transgenic; promoter.

XX Gallus gallus.

XX US200317516-A1.

XX 18-SEP-2003.

XX 14-MAR-2002; 2002US-00099663.

XX 14-MAR-2002; 2002US-00099663.

XX (HORS/) HORSEMAN N D.

XX (PRAT/) PRATT S L.

XX Horseman ND, Pratt SL;

XX WPI; 2003-898653/82.

PT New nucleic acid molecule comprising an isolated avian gut-specific gene
PT expression control region, useful for regulating heterologous nucleic
PT acids in transgenic avians, and for generating transgenic birds.

PS Claim 1; SEQ ID NO 2; 28bp; English.

CC The invention relates to an isolated nucleic acid comprising an isolated
CC avian gut-specific gene expression control region appearing as
CC ADL90127(Chicken intestinal fatty acid binding protein, IFABP, gene, 5'
CC region or ADL90128 (Chicken IFABP promoter) or its degenerate variant.
CC Also included are a recombinant DNA molecule comprising an isolate avian
CC gut-specific gene expression control region operably linked to a nucleic
CC acid insert encoding a polypeptide, an expression vector that integrates
CC into a host cell (and comprising the isolated avian gut-specific gene
CC expression control region), expressing a heterologous polypeptide in a
CC host cell (by transfecting a eukaryotic cell with the recombinant DNA
CC molecule, and culturing the transfected cell in a medium suitable for
CC expression of a heterologous polypeptide under the control of an avian
CC intestinal fatty acid binding protein (IFABP) or cp35 gene expression
CC control region encoded by the recombinant DNA molecule), a eukaryotic
CC cell transformed with the expression vector (or its progeny, which
CC expresses a heterologous polypeptide) and a transgenic avian having a
CC heterologous polynucleotide sequence comprising the nucleic acid insert.
CC The nucleic acids are useful for regulating heterologous nucleic acids in
CC transgenic avians, as probes in nucleic acid hybridisation assays for
CC detecting the IFABP gene expression control region, and for generating
CC transgenic birds. The present sequence is the Chicken intestinal fatty
CC acid binding protein, IFABP, gene, promoter.

XX SQ Sequence 336 BP; 116 A; 62 C; 63 G; 95 T; 0 U; 0 Other;

Query Match 63.7%; Score 326; DB 11; Length 336;
Best Local Similarity 100.0%; Pred. No. 7.9e-79;

Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATTATTTATTTTTCATTAGATAGCCGGTTTTTACTCAACTCAATAATAGATGAACGAATG 246
DB 1 ATTATTTATTTTTCATTAGATAGCCGGTTTTTACTCAACTCAATAATAGATGAACGAATG 60
QY 247 AATGGTTAGTACGTTTATTAAGAAGAGTAAATGAATCTATCATCATTTGAGGCA 306
DB 61 AATGGTTAGTACGTTTATTAAGAAGAGTAAATGAATCTATCATCATTTGAGGCA 120
QY 307 TAAGGAGGAGAGATTCAGCAACAAGTGTGCTTCAAGTGAAGAAACAAGTTAACTAAA 366
DB 121 TAAGGAGGAGAGATTCAGCAACAAGTGTGCTTCAAGTGAAGAAACAAGTTAACTAAA 180
QY 367 GTGACCCCTCTCTTGAACAAGTCAATGCCACAGTTGAGCTTTAGCCAGCACAATCATCA 426
DB 181 GTGACCCCTCTCTTGAACAAGTCAATGCCACAGTTGAGCTTTAGCCAGCACAATCATCA 240
QY 427 TGTAAATGCTTCCGTGATTAAGCCTGTTCAATTAATTCCTTTGCAAGCTGTGACTTA 486
DB 241 TGTAAATGCTTCCGTGATTAAGCCTGTTCAATTAATTCCTTTGCAAGCTGTGACTTA 300
QY 487 CCAGAAAGTGTGCTTACAGACAGAAAG 512

Query Match 8.3%; Score 42.4; DB 4; Length 5908;
Best Local Similarity 54.5%; Pred. No. 0.46; Indels 0; Gaps 0;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 138 TATTAAGTCCTATGCAACAAAGAAATGCAATACATCTCTTACTGTCATTTATTTT 197
|||||
DB 4703 TATTAACCTCTTTTATATAAAACAAATATACATATATCTCAATAAACCATCTTTTAA 4644
|||||

QY 198 CATTAGATAGCCGCTTTTCTACACCTCAATTAAGTAAAGCAAGATGAATGGTTAGT 257
|||||
DB 4643 TATTAATATTTTATTTTACTACCTCTTCTCAAAATTTAAAAATTAATCTATACCTTAAT 4584
|||||

QY 258 GACTGTTTATAAAGAGATTAATTAAGATCTATCA 293
|||||
DB 4583 TCCCTTTATCCAAATTAATTAATCTATATACGACA 4548
|||||

RESULT 5
ABK28232/c
ID ABK28232 standard; DNA; 5908 BP.
XX
AC ABK28232;
XX
DT 23-APR-2002 (first entry)
XX
DE DNA transcription associated complementary genomic DNA #53.
XX
XX DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
XX PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
XX single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
XX viral infection; Sezary syndrome; haematological disorder; tuberculosis;
XX immunological disorder; Werner syndrome; developmental disorder;
XX psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
XX neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
XX myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
XX angiodysplasia; congenital heart disease; HDR syndrome; gene therapy;
XX polyglutamine disorder; solid tumour.
XX
OS Unidentified.
XX
XX WO200192565-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 06-APR-2001; 2001WO-EP003973.
XX
XX PR 06-APR-2000; 2000DE-01019058.
XX PR 07-APR-2000; 2000DE-01019173.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX
XX PA (EPIC-) EPIGENOMICS AG.
XX
XX PI Olek A, Piepenbrock C, Berlin K;
XX
XX DR MPI; 2002-090046/12.
XX
XX PT New nucleic acids or oligomers, useful for diagnosing or treating
XX diseases associated with DNA transcription, e.g. immunological disorders,
XX Werner syndrome, psoriasis, myocardial infarction, solid tumors or
XX cancer.
XX
XX PS Claim 1; SEQ ID NO 106; 32pp; English.
XX
XX CC The invention relates to a nucleic acid, which comprises a segment of the
XX chemically pretreated DNA of genes associated with DNA transcription from
XX one of 346 sequences, and an oligomer, in particular an oligonucleotide
XX or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
XX to the chemically pretreated DNA of genes associated with DNA
XX transcription. The set of oligomer probes are useful for detecting the
XX cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
XX in a chemically pretreated genomic DNA. The nucleic acids are useful for

CC diagnosing or treating diseases associated with DNA transcription
CC (particularly with the methylation status) e.g. adenosine deaminase
CC deficiency, viral infection, retroviral infection, Sezary syndrome,
CC haematological disorders, immunological disorders, Werner syndrome,
CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
CC neurological disorders, neurodegenerative disorders, Waardenburg
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
CC associated genomic DNA molecules of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification but
CC was obtained in electronic format directly from the European Patent
CC Office

SQ Sequence 5908 BP; 1706 A; 149 C; 1421 G; 2632 T; 0 U; 0 Other;

Query Match 8.3%; Score 42.4; DB 6; Length 5908;
Best Local Similarity 54.5%; Pred. No. 0.46;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 138 TATTAAGTCCTATGCAACAAAGAAATGCAATACATCTCTTACTGTCATTTATTTT 197
|||||
DB 4703 TATTAACCTCTTTTATATAAAACAAATATACATATATCTCAATAAACCATCTTTTAA 4644
|||||

QY 198 CATTAGATAGCCGCTTTTCTACACCTCAATTAAGTAAAGCAAGATGAATGGTTAGT 257
|||||
DB 4643 TATTAATATTTTATTTTACTACCTCTTCTCAAAATTTAAAAATTAATCTATACCTTAAT 4584
|||||

QY 258 GACTGTTTATAAAGAGATTAATTAAGATCTATCA 293
|||||
DB 4583 TCCCTTTATCCAAATTAATTAATCTATATACGACA 4548
|||||

RESULT 6
ABL70313/c
ID ABL70313 standard; DNA; 10048 BP.
XX
XX ABL70313;
XX
XX DT 01-JUN-2002 (first entry)
XX
XX DE Chemically treated cell signalling DNA sequence#102.
XX
XX XX Cell signalling; cytosine methylation; cell signalling disease; cancer;
XX tumour; cytosine; ds.
XX
XX OS Unidentified.
XX
XX XX WO200202807-A2.
XX
XX XX 10-JAN-2002.
XX
XX XX 29-JUN-2001; 2001WO-EP007471.
XX
XX XX 30-JUN-2000; 2000DE-01032529.
XX
XX XX 01-SEP-2000; 2000DE-01043826.
XX
XX PA (EPIC-) EPIGENOMICS AG.
XX
XX PI Olek A, Piepenbrock C, Berlin K;
XX
XX DR MPI; 2002-154758/20.
XX
XX XX Nucleic acid, useful for diagnosis and therapy of diseases associated
XX with cell signalling e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with cell signalling.
XX
XX PS Claim 1; SEQ ID NO 203; 24pp + Sequence Listing; English.
XX
XX CC The invention relates to a nucleic acid comprising a sequence of at least
XX 18 bases of a segment of chemically pretreated DNA of genes associated
XX with cell signalling. The activity of the modified sequences of the

CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling. Note: The
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office

XX
SQ Sequence 10048 BP; 2873 A; 57 C; 1888 G; 5225 T; 0 U; 5 Other;
Query Match 8.2%; Score 41.8; DB 6; Length 10048;
Best Local Similarity 46.4%; Pred. No. 0.81;
Matches 136; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 100 ATACGTTTCTCCAGTTGTAAGACCTCCACCAAGTAAAGTCTATGCAACAAG 159
DB 5471 ATATTATATCTCTAAACTAAATTCCTCTACAAAAAAATATATATCTATAC 5412
QY 160 AAAATGCAATACATCTCTAGTCTCATATATTTTCTATGATAGCCGTTTTC 219
DB 5411 TCCAAATCCAAAAATCTATTTTAAACCTATTAATCCCTTACAAATACCTTCAC 5352
QY 220 TACAATCAAAATAGATGAAGATGAATGAGTGTAGTACTGTTATTAAGAAGTAA 279
DB 5351 AATAACTAATATAAAAACACACATTAATTCCTTAATCTAATTCAAAATACAAAAAC 5292
QY 280 TAAAGTACTATCATCATTTAGGCAATTAAGGAGGAGAGATTCGCAAAAGTGTCT 339
DB 5291 TATTAATATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5232
QY 340 TCAAGTGAAGAAACAAGTTAAAGTAAAGGAGGAGAGATTCGCAAAAGTGTCT 392
DB 5231 AAACAAACAAACAAACAAACAAACAAACAAACAAATTCAAACCTATATCTA 5179

RESULT 7
AAS61251/c
ID AAS61251 standard; DNA; 10048 BP.
XX
XX AAS61251;
DT 29-JAN-2002 (first entry)
XX
DE Human gene regulation-associated gene oligonucleotide #206.
XX
XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
XX cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
XX asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
XX renal disease; Preeclampsia; cardiac allograft vascular disease;
XX colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
XX immunostimulant; cardiac; anti-inflammatory; coagulant; antiaesthetic;
XX nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
XX
OS Homo sapiens.
XX
XX MO200177375-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001MO-EP003968.
XX
PR 06-APR-2000; 2000DE-EP003968.
XX
PR 07-APR-2000; 2000DE-01019058.
XX
PR 30-JUN-2000; 2000DE-01019173.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIC-) EPIGENOMICS AG.
PA

XX
PI Olek A, Piepenbrock C, Berlin K;
DR WPI; 2002-017470/02.
XX
XX New nucleic acid sequences from chemically modified genes associated with
PT gene regulation, useful for analyzing cytosine methylations for diagnosis
XX and therapy of diseases e.g. severe combined immunodeficiency disease.
PS Claim 1; SEQ ID NO 211; 26bp; English.
XX
XX The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The chemical
CC pretreatment converts cytosine bases unmethylated at the 5-position to
CC uracil or another base with hybridisation behaviour dissimilar to
CC cytosine, to enable analysis of cytosine methylations. The DNA sequences,
CC oligomers (or sets/arrays) and method are useful in the diagnosis of
CC diseases (or predisposition to diseases) associated with gene regulation
CC and in therapy of such diseases, by enabling analysis of the cytosine
CC methylation patterns of such genes. Kits are provided. They are
CC especially useful in diagnosis and therapy of e.g. severe combined
CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
CC syndrome, renal disease, preeclampsia, graft versus-host disease. The
CC present sequence is a sequence included in the sequence data for this
CC specification and is associated with the human gene regulation-associated
CC genes. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO.int/ftp/pub/published_pct_sequences

XX
SQ Sequence 10048 BP; 2873 A; 57 C; 1888 G; 5225 T; 0 U; 5 Other;
Query Match 8.2%; Score 41.8; DB 6; Length 10048;
Best Local Similarity 46.4%; Pred. No. 0.81;
Matches 136; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 100 ATACGTTTCTCCAGTTGTAAGACCTCCACCAAGTAAAGTCTATGCAACAAG 159
DB 5471 ATATTATATCTCTAAACTAAATTCCTCTACAAAAAAATATATATCTATAC 5412
QY 160 AAAATGCAATACATCTCTAGTCTCATATATTTTCTATGATAGCCGTTTTC 219
DB 5411 TCCAAATCCAAAAATCTATTTTAAACCTATTAATCCCTTACAAATACCTTCAC 5352
QY 220 TACAATCAAAATAGATGAAGATGAATGAGTGTAGTACTGTTATTAAGAAGTAA 279
DB 5351 AATAACTAATATAAAAACACACATTAATTCCTTAATCTAATTCAAAATACAAAAAC 5292
QY 280 TAAAGTACTATCATCATTTAGGCAATTAAGGAGGAGAGATTCGCAAAAGTGTCT 339
DB 5291 TATTAATATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5232
QY 340 TCAAGTGAAGAAACAAGTTAAAGTAAAGGAGGAGAGATTCGCAAAAGTGTCT 392
DB 5231 AAACAAACAAACAAACAAACAAACAAACAAACAAATTCAAACCTATATCTA 5179

RESULT 8
AAS46439/c
ID AAS46439 standard; DNA; 7040 BP.
XX
XX AAS46439;
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #161.
XX
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
XX tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX cytosine methylation; ds.
XX
OS Homo sapiens.
XX

```

XX XX WO200168912-A2.
XX PN
XX PD 20-SEP-2001.
XX PF 15-MAR-2001; 2001WO-BP002955.
XX PR 15-MAR-2000; 2000DE-01013847.
XX PR 06-APR-2000; 2000DE-01019058.
XX PR 07-APR-2000; 2000DE-01019173.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI, 2001-602752/68.
XX PT
XX PR Fragments of chemically modified genes associated with tumor suppressor
XX PT genes and oncogenes, useful in designing primers and probes for analyzing
XX PT diseases associated with cytosine methylation state e.g. cancer.
XX PS Claim 1; SEQ ID NO 161; 27pp; English.
XX XX
XX CC The invention relates to a nucleic acid comprising a sequence of 18
XX CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX CC bisulphite, of genes associated with tumour suppression and oncogenes
XX CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
XX CC 500 are missing from the sequence listing) sequences (Ss) and sequences
XX CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
XX CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
XX CC probes for detecting the cytosine methylation state and/or single
XX CC nucleotide polymorphisms and also to be used in an array for analysing
XX CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
XX CC probes can also be used in a method for ascertaining genetic and/or
XX CC epigenetic parameters for the diagnosis and/or therapy of existing
XX CC diseases or the predisposition to specific diseases, by analysing
XX CC cytosine methylations. The parameters may be compared to another set of
XX CC genetic and/or epigenetic parameters, the differences serving as basis
XX CC for diagnosis and/or prognosis events which are disadvantageous to
XX CC patients. The present sequence is one of the 533 genomic sequences
XX CC derived from tumour suppressor genes and oncogenes. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;
XX
XX Query Match 7.9%; Score 40.6; DB 4; Length 7040;
XX Best Local Similarity 50.8%; Pred. No. 1.5;
XX Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
XX
XX QY 101 TAGCGTTTTCGAGTTGTAATAAGACCTCCACAGATTAAGTCTATGACAAAGA 160
XX DB 3586 TATATTTTTAAAAAACAATTAATAAATTCCTTTCTAAACACTATATATTAATTT 3527
XX
XX QY 161 AAATGCAATACATCTCTTAGTCATATATATTTTCATTAGATAGCGGTTTTTACT 220
XX DB 3526 ATATATTAACCAATTTTAATTAATAACAATTCATATTTTAAATATATTAACCTTTAAAAA 3467
XX
XX QY 221 ACAACTCAAAATAGATGAACAGAAATGAATGGGTAGTGACTGTTTATTAAGAAGATAT 280
XX DB 3466 ATTAATTAATAAACAATAAACCAATTTACAAAACAATTAATAATCATATAAACCCATAAAT 3407
XX
XX QY 281 AAAGATACAT 291
XX DB 3406 AAAATTTCTAT 3396
XX
XX RESULT 9
XX ABK33963/c
XX ID ABK33963 standard; DNA; 7040 BP.

```

ABK33963;
18-JUN-2002 (first entry)
Human DNA for staging of Astrocytomas #24.
Human, ds; astrocytoma, cytosstatic; staging; cysteine methylation; CpG; bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry; matrix assisted laser desorption/ionization mass spectrometry.
Homo sapiens.
WO200202808-A2.
10-JAN-2002.
02-JUL-2001; 2001WO-EP007538.
30-JUN-2000; 2000DE-01032529.
01-SEP-2000; 2000DE-01043826.
(EPig-) EPIGENOMICS AG.
Olek A, Piepenbrock C, Berlin K;
WPI; 2002-171649/22.
Novel chemically modified genomic DNA sequences, useful in the characterization, classification, differentiation, grading, staging, treatment and/or diagnosis of astrocytomas or predisposition to astrocytomas.
Claim 1; SEQ ID NO 47; 37pp: English.
The invention relates to a nucleic acid comprising a sequence (1) of at least 18 bases in length of a segment of chemically pre-treated genomic DNA which has any one of the sequences of (ABK33919-ABK34032) or its complement. Also included are an oligonucleotide or peptide nucleic acid (or set thereof) of at least 9 nucleotides which hybridises to (1), primers for (1), probes for detecting cytosine methylation or single-nucleotide polymorphisms (SNP) in (1), an array of oligomers or peptide nucleic acids for analysing diseases associated with the methylation states of the CpG dinucleotides of (1). The array is useful for determining genetic and/or epigenetic parameters, classification, differentiation, grading, staging, treatment and/or diagnosis of astrocytomas, or the predisposition to astrocytomas by analysing cytosine methylations, involves obtaining a biological sample containing genomic DNA, extracting the genomic DNA, converting cytosine bases which are unmethylated at the 5-position, in the genomic DNA sample, to uracil or another base which is dissimilar to cytosine in terms of hybridisation behaviour, by chemical treatment and amplifying chemically pre-treated genomic DNA fragments using the array and a polymerase, where the amplificates carry a detectable label. The method further involves identifying methylation status of one or more cytosine positions, and analysing methylation status of the cytosine positions by reference to one or more data sets. The genomic DNA is chemically treated by using a bisulphite, hydrogen sulphite or disulphite. The amplification step amplifies DNA which is of particular interest in astrocytoma or brain tissue, based on the specific genomic methylation status of brain tissues, as opposed to background DNA. The amplificates carry a fluorescent label or radionuclide. Optionally, the labels of the amplificates are detachable molecule fragments having a typical mass which are detected in a mass spectrometer. The fragments of chemically pre-treated genomic DNA to be amplified, have a single positive or negative charge for a better detectability in the mass spectrometer. Preferably, the amplificates or fragments of the amplificates are detected by matrix assisted laser desorption/ionization mass spectrometry (MALDI) or using electron spray mass spectrometry (ESI). The present sequence is one of the chemically pre-treated reference DNA samples of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences

XX Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;
SQ
Query Match 7.9%; Score 40.6; DB 6; Length 7040;
Best Local Similarity 50.8%; Pred. No. 1.5;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 101 TACGTTTCCGATGTAATGAAGCCGCCAGTAAAGTCTTATGCAACAAGA 160
DB 3586 TATATTTTAAAAAATCAATTAATTAATTCCTTTCTTAAACACTTATATTAAT 3527
QY 161 AATGTCATACATTCCTTAGTCATTAATTTTCAATAGATAGCCGGTTTTTACT 220
DB 3526 ATATATACACATTTTAAATTAATACATTTCTATTTTAAATATTAATTAATTA 3467
QY 221 ACACCTCAATTAAGATGACAGATGAATGAGGTAGTACTGTTTAAAGAAAGTAAT 280
DB 3466 ATTAATTAATTAATCAATTAACATTTACAAAACAAATTAATCTATTAATTAAT 3407
QY 281 AAGATACCTAT 291
DB 3406 AATATTTCTAT 3396
RESULT 10
ABZ10179/c
ID ABZ10179 standard; DNA; 7040 BP.
XX
XX ABZ10179;
XX
XX 16-JAN-2003 (first entry)
XX
XX Haematopoietic cell proliferation disorder related DNA sequence #319.
XX
XX Human; haematopoietic cell proliferation disorder; cytostatic;
XX gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
XX cytosine methylation state; gene; ds.
XX
XX Homo sapiens.
XX
XX WO20027272-A2.
XX
XX 03-OCT-2002.
XX
XX 26-MAR-2002; 2002WO-EP003401.
XX
XX 26-MAR-2001; 2001US-0278333P.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Berlin K, Braun A, Dietler J, Guetig D, Howe A, Mueller J;
XX Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
XX Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
XX Schwobe I, Ziebarth H;
XX WPI; 2003-018942/01.
XX
XX Detecting and differentiating between hematopoietic cell proliferative
XX disorders, comprises contacting a target nucleic acid with a reagent that
XX distinguishes between methylated and non-methylated CpG dinucleotides.
XX
XX Claim 28; SEQ ID NO 319, 117p; English.
XX
XX The present invention describes a method for detecting and
XX differentiating between haematopoietic cell proliferative disorders
XX associated with at least 1 gene and/or their regulatory regions in a
XX subject. The method comprises contacting a target nucleic acid in a
XX biological sample obtained from the subject with at least 1 reagent,
XX which distinguishes between methylated and non-methylated CpG
XX dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
XX represent specifically claimed nucleotide sequences from the present
XX invention. Oligonucleotides from the present invention can be used; for
XX differentiating between healthy haematopoietic cells and proliferative

CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
SQ Sequence 7040 BP; 2018 A; 0 C; 1566 G; 3456 T; 0 U; 0 Other;
Query Match 7.9%; Score 40.6; DB 8; Length 7040;
Best Local Similarity 50.8%; Pred. No. 1.5;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 101 TACGTTTCCGATGTAATGAAGCCGCCAGTAAAGTCTTATGCAACAAGA 160
DB 3586 TATATTTTAAAAAATCAATTAATTAATTCCTTTCTTAAACACTTATATTAAT 3527
QY 161 AATGTCATACATTCCTTAGTCATTAATTTTCAATAGATAGCCGGTTTTTACT 220
DB 3526 ATATATACACATTTTAAATTAATACATTTCTATTTTAAATATTAATTAATTA 3467
QY 221 ACACCTCAATTAAGATGACAGATGAATGAGGTAGTACTGTTTAAAGAAAGTAAT 280
DB 3466 ATTAATTAATTAATCAATTAACATTTTACAAAACAAATTAATCTATTAATTAAT 3407
QY 281 AAGATACCTAT 291
DB 3406 AATATTTCTAT 3396
RESULT 11
ABZ10033/c
ID ABZ10033 standard; DNA; 7040 BP.
XX
XX ABZ10033;
XX
XX 16-JAN-2003 (first entry)
XX
XX Haematopoietic cell proliferation disorder related DNA sequence #173.
XX
XX Human; haematopoietic cell proliferation disorder; cytostatic;
XX gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
XX cytosine methylation state; gene; ds.
XX
XX Homo sapiens.
XX
XX WO20027272-A2.
XX
XX 03-OCT-2002.
XX
XX 26-MAR-2002; 2002WO-EP003401.
XX
XX 26-MAR-2001; 2001US-0278333P.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Berlin K, Braun A, Dietler J, Guetig D, Howe A, Mueller J;
XX Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
XX Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
XX Schwobe I, Ziebarth H;
XX WPI; 2003-018942/01.
XX
XX Detecting and differentiating between hematopoietic cell proliferative
XX disorders, comprises contacting a target nucleic acid with a reagent that
XX distinguishes between methylated and non-methylated CpG dinucleotides.
XX

PS	Claim 28; SEQ ID NO 173; 117pp; English.
XX	
CC	The present invention describes a method for detecting and
CC	differentiating between haematopoietic cell proliferative disorders
CC	associated with at least 1 gene and/or their regulatory regions in a
CC	subject. The method comprises contacting a target nucleic acid in a
CC	biological sample obtained from the subject with at least 1 reagent,
CC	which distinguishes between methylated and non-methylated CpG
CC	dinucleotides within the target nucleic acid. AB209861 to AB21118
CC	represent specifically claimed nucleotide sequences from the present
CC	invention. Oligonucleotides from the present invention can be used for
CC	differentiating between healthy haematopoietic cells and proliferative
CC	disorder haematopoietic cells; for differentiating between acute
CC	lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC	determining the cytosine methylation state and/or single nucleotide
CC	polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC	related sequences and their complements; and as primers for the
CC	amplification of haematopoietic cell proliferation disorder related DNA
CC	sequences. The nucleotide sequences from the present invention can also
CC	be used for detecting a predisposition to, differentiation between
CC	subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC	haematopoietic cell proliferative disorders. The present method enables a
CC	highly specific classification of haematopoietic cell proliferative
CC	disorders allowing for improved and informed treatment of patients
XX	
XX	Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;
XX	
Query Match	7.9%; Score 40.6; DB 8; Length 7040;
Best Local Similarity	50.8%; Pred. No. 1.5;
Matches	97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
OY	101 TACGTTTCTCCAGTGTATTAAGACCTCCACAGATTAAGTCTATGACAAAGA 160
DB	3586 TATATTTTAAAAAATCAATATATAAATTCCTTCTTAAAAACCTATATATTAATTT 3522
OY	161 AAATGTCAATACATCTCTTAGTGTCCATATATTTTCAATAGATGACGGTTTACT 220
DB	3526 ATATATTAACCATTTTAAATATTAACAATTCATTTTAAATATATTAACCTTTAAAAA 3467
OY	221 ACAACTCAAAATGAAGTGAACAGAAATGAATGGTGTAGTGA CTGTTTATTAAGAAGATAT 280
DB	3466 ATAAATTAATATAACATAAACCTTTACAAAAACAATTTAAATCTATTAACCCATAAAT 3407
OY	281 AAAGTACTAT 291
DB	3406 AAAATTTCTAT 3396
XX	
RESULT 12	
ID	ADA20348/C
XX	ADA20348 standard; DNA; 7040 BP.
XX	
AC	ADA20348;
XX	
XX	20-NOV-2003 (first entry)
DE	Prostate tumour related genomic DNA sample #7.
XX	
XX	cytostatic; gene therapy; genetic marker; epigenetic parameter;
KW	classification; differentiation; diagnosis; prostate tumour;
KW	prostate cancer; cytosine methylation; uracil;
KW	single nucleotide polymorphism; SNP; prostate carcinoma; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO2002103042-A2.
XX	
PD	27-DEC-2002.
XX	
PF	14-JUN-2002; 2002WO-BP006605.
XX	
FR	14-JUN-2001; 2001DE-01028508.
XX	

Query Match	Best Local Match	Similarity	Score	DB	Length	7040;
Matches	97;	Conservative	50.88;	Pred. No.	1.5;	Mismatches
				94;	Indels	0;
					Gaps	0;
101	TACGTTTTCCTCCAGTTGTATTAAGACCCCTCCACAGATTAAGTCCCTATGCAACAAGA	160				
3586	TATATTTTAAATAATCAATTAATAAATTCCTCTCTAATAACACTTATATTAATTT	3527				
161	AAATGCAATACATTCTCTTAGTCTCATTTATTTTCAATTAGATAGCCGCTTTTAACT	220				
3526	ATAATATACCCACATTTTAAATATTAACAACTTATTTTAAATATATTAACCTTTAAAAA	3467				
221	ACAACCTAAATTAAGATGAACGATGATGGGTTAGTGCCTGTTTAAAGAAGATTAAT	280				
3466	ATTAATTTATTAATTAACATTAACCATTTACAAAAACAATTTAAATCTATTAACCACTTAAT	3407				
281	AAAGATACTAT	291				
3406	AAAATTTCTAT	3396				
RESULT 13	ADA84155/C	ADA84155 standard; DNA; 7040 BP.				
AC	ADA84155;					
DT	20-NOV-2003	(first entry)				
DE	Human renal/prostate carcinoma associated DNA SEQ ID NO:13.					
KM	ds; renal cancer; prostate cancer; cytosine methylation; single nucleotide polymorphism; histological; cytological.					
OS	Homo sapiens.					
PN	WO2002103041-A2.					
PD	27-DEC-2002.					
PF	14-JUN-2002; 2002WO-EP006603.					
PR	14-JUN-2001; 2001DE-01028509.					

XX

RESULT 15
ADE84115/C

ID ADE84115 standard; DNA; 7040 BP.
XX
AC ADE84115;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human lymphoid cell proliferative disorder gene derived DNA #51.
XX
KM de; lymphoid cell proliferative disorder; methylation;
KM methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
KM diffuse large B-cell lymphoma; mantle cell lymphoma;
KM chronic lymphocytic leukemia; small lymphocytic lymphoma;
KM follicular lymphoma; diagnosis; prognosis.
XX
OS Homo sapiens.
XX
PN MO2003044226-A2.
XX
PD 30-MAY-2003.
XX
PF 25-NOV-2002; 2002MO-EP013265.
XX
PR 23-NOV-2001; 2001DE-01057491.
PR 28-DEC-2001; 2001DE-01064501.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;
DR WPI; 2003-457621/43.
XX
XX
PT Detecting and differentiating between lymphoid cell proliferative
PT disorders comprises contacting a target nucleic acid with at least one
PT reagent that distinguishes between methylated and non-methylated CpG
PT dinucleotides.
XX
PS Claim 26; SEQ ID NO 111; 448bp; English.
XX
XX
CC The invention relates to a method of detecting and differentiating
CC between lymphoid cell proliferative disorders associated with at least
CC one gene and/or their regulatory regions in a subject by contacting a
CC target nucleic acid in a biological sample obtained from the subject with
CC at least one reagent or series of reagents that distinguish between
CC methylated and non-methylated CpG dinucleotides within the target nucleic
CC acid. The genes and/or their regulatory regions are preferably selected
CC from MDRI, CSNK2B, EGR4, AR, CDK4, RB2, CDC25A, GPRbeta, MYOD1, CDH3,
CC MYCL1, EIK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN1B, CDKN2A, CDKN2B, FOS,
CC GSTP1, HIC-1, MGMT, MLH1, MDS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C,
CC GSK3beta, ESRI, APAF1, BAK1, BAX or HOXA5. Oligomers, peptide nucleic
CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences
CC of the genes are useful for detecting the methylation state of all the
CC CpG dinucleotides within one or more the sequences, or their complements,
CC for determining the cytosine methylation state and or single nucleotide
CC polymorphisms (SNPs), and for differentiating at least two of the medical
CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,
CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular
CC lymphoma. They are also useful for detecting of a predisposition to,
CC differentiation between subclases, diagnosis, prognosis, treating and/or
CC monitoring of lymphoid cell proliferative disorder. This sequence
CC represents a nucleic acid of a pretreated genomic DNA derived from the
CC above mentioned genes.
XX
SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 U; 0 Other;

Query Match 7.9%; Score 40.6; DB 10; Length 7040;
Best Local Similarity 50.8%; Pred. No. 1.5;
Matches 97; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 101 TACGTTTCCCGTGTATTAAGACCCGCCACGATATAAGTCTATGCAACAAAGA 160
DB 3586 TATATTTTAAATAAATCAAAATTAATTCCTTTCTTAATAACACTTATATTAATTT 3527
QY 161 AATGTCATACATTCCTTAGTCTCATTTATTTTATTTTATTTAGATAGCCGGTTTAACT 220

DB 3526 ATATATAACCAATTTTATATATAACAAATCTATTTTAAATATATTAACCTTAAAAAA 3467
QY 221 ACAACTCAATTAAGATGAACGAATGAATGGGTAGTGAATCTTTATATAAGAAAGTAAAT 280
DB 3466 ATTAATTTAATAAACAATAACCAATTTTACAAAAACAATTTTAAATCTTAAACCTTAAAT 3407
QY 281 AAAGATACCTAT 291
DB 3406 AAAATTTCTAT 3396

Search completed: November 21, 2005, 05:53:43
Job time : 410.736 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 05:12:21 ; Search time 82.8113 Seconds
(without alignments)
6639.056 Million cell updates/sec

Title: US-10-099-663-2

Perfect score: 336
Sequence: 1 attatatttcttcattagata.....agacagaagaatgcatctta 336

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgm2_6/prodata/c/ina/5A_COMB.seq:*
2: /cgm2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgm2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgm2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgm2_6/prodata/1/ina/PCBUS_COMB.seq:*
6: /cgm2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37.2	11.1	1141	4 US-09-806-708B-22	Sequence 22, Appl
2	36.8	11.0	43117	4 US-09-949-016-17589	Sequence 17589, A
3	36.2	10.8	1587	4 US-09-710-279-1929	Sequence 1929, Ap
4	36.2	10.8	1878	3 US-09-134-001C-1667	Sequence 1667, Ap
5	36.2	10.8	3000	4 US-09-710-279-3907	Sequence 3907, Ap
6	36.2	10.8	3368	4 US-09-710-279-4169	Sequence 4169, Ap
7	36	10.7	168394	4 US-09-949-016-13002	Sequence 13002, A
8	35.8	10.7	7218	1 US-08-232-463-14	Sequence 14, Appl
9	35.6	10.6	601	4 US-09-949-016-44395	Sequence 44395, A
10	35.6	10.6	601	4 US-09-949-016-58898	Sequence 58898, A
11	35.6	10.6	139562	4 US-09-949-016-13451	Sequence 13451, A
12	35.6	10.6	256171	4 US-09-949-016-12822	Sequence 12822, A
13	35.6	10.6	256176	4 US-09-949-016-15524	Sequence 15524, A
14	35.2	10.5	601	4 US-09-949-016-134974	Sequence 134974, A
15	35.2	10.5	663	3 US-08-998-416-187	Sequence 187, App
16	35.2	10.5	666	3 US-08-998-416-779	Sequence 779, App
17	35.2	10.5	719	3 US-08-998-416-1138	Sequence 1138, Ap
18	35.2	10.5	856	3 US-08-998-416-289	Sequence 289, App
19	35	10.4	399	4 US-09-621-976-8976	Sequence 8976, Ap
20	34.8	10.4	78125	4 US-09-949-016-16006	Sequence 16006, A
21	34.6	10.3	601	4 US-09-949-016-173718	Sequence 173718, A
22	34.6	10.3	601	4 US-09-949-016-173765	Sequence 173765, A
23	34.6	10.3	81819	4 US-09-949-016-16661	Sequence 16661, A
24	34.6	10.3	81819	4 US-09-949-016-16661	Sequence 16661, A
25	34.4	10.2	1137	4 US-09-107-532A-2846	Sequence 2846, Ap
26	34.4	10.2	1664976	4 US-08-916-421B-1	Sequence 1, Appl
27	34.4	10.2	1664976	4 US-09-692-570-1	Sequence 1, Appl

28	34.2	10.2	98844	3 US-09-791-211-10	Sequence 10, Appl
29	34.2	10.2	143776	4 US-09-949-001-29	Sequence 29, Appl
30	34.2	10.2	144034	4 US-09-949-001-35	Sequence 35, Appl
31	33.6	10.0	2119	3 US-09-240-639-7	Sequence 7, Appl
32	33.6	10.0	2119	4 US-09-908-510A-7	Sequence 7, Appl
33	33.6	10.0	2119	4 US-09-905-744B-7	Sequence 7, Appl
34	33.6	10.0	2119	4 US-10-107-660-7	Sequence 7, Appl
35	33.6	10.0	2119	4 US-10-107-576-7	Sequence 7, Appl
36	33.6	10.0	2119	4 US-09-905-732B-7	Sequence 7, Appl
37	33.6	10.0	2119	4 US-09-905-743B-7	Sequence 7, Appl
38	33.4	9.9	31385	4 US-09-949-016-13310	Sequence 13310, A
39	33.4	9.9	73788	4 US-09-949-016-12358	Sequence 12358, A
40	33	9.8	2642	1 US-08-178-242-4	Sequence 4, Appl
41	33	9.8	2642	2 US-08-955-091-4	Sequence 4, Appl
42	33	9.8	2642	3 US-09-225-510-4	Sequence 4, Appl
43	33	9.8	2870	1 US-08-178-242-14	Sequence 14, Appl
44	33	9.8	2870	2 US-08-955-091-14	Sequence 14, Appl
45	33	9.8	2870	3 US-09-225-510-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-806-708B-22
Sequence 22, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806, 708B
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147, 133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1141)
OTHER INFORMATION: consensus sequence of A.c., I.a., and B.n. FAEl promoters
US-09-806-708B-22

Query Match	Query Match	Score	DB 4	Length	Indels	Gaps
Best Local Similarity	11.1%	10.7%	Pred. No. 0.24	1141	0	0
Matches	31	Conservative 123	Mismatches 135	Indels 0	Gaps 0	
QY	1	ATTATATTCTTCTAGATAGCGGTTTCTTCTCACTCAATTAATGATGAACAGATG	60			
DB	285	ATACMCRAATYWGNNABVSTCTTWSKTTXRTSCWANNCAAGDANKDKHKKWSAAWG	344			
QY	61	AATGGTTATGACTGTGTTATTAAGAAGATATAAAGTACTATCATATTGAGCAAA	120			
DB	345	VYNNNNNNNNWYTKAHHBARDWVHSAKKHANAHAHSRKKWBYRKTIVNNNGT	404			
QY	121	TAAAGGAGGAGAGATTCAAGCAACGTGCTTCAAGTGAAGAAAACAAGTTAAACTAA	180			
DB	405	TWKKRWAMWYKMDMDWBGTYNNNNNGRTYGTWKKWYTKKANNCKRARDHKT	464			
QY	181	GTGACCCCTCTCTGCAAGATCAATGACAGTTGAGCTTAAAGCAGCATCATCATCA	240			
DB	465	CTHNNNTTWKMTKTYNNNCYKSMTGKSHRBAAYTYWMMWRRAHNNNNMDYWMKA	524			
QY	241	TGTAAATTCCTTCTCGATTAACCTGTTATTAATTCCTTTGCAAGC	289			
DB	525	CTWYKBYCCKWNNWYAAWYTKSSWNYTSRYYRKKTNNSWRWSDTRSM	573			

US-09-949-016-17589/C
; Sequence 17589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17589
; LENGTH: 43117
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17589

Query Match 11.0%; Score 36.8; DB 4; Length 43117;
Best Local Similarity 52.6%; Pred. No. 1.3;
Matches 80; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 30 TTACTCAACTCAATTAAGTGAACAGATGAATGGCTTGTGACTGTTTATTAAGAGA 89
DB 20807 TGAACCAATATATGATGTGTAATAATTATGGAGGATTAATTAAGTATGAATGCA 20748
QY 90 GTAATAAGTACTATCATCATTTGAGGCAATTAAGGAGGAGAGATTCAGCAAAACGTG 149
DB 20747 CCTTAAGAGCAAAATGTAATGTAAGGAATAGCCTCTGGCCGGAATCTGAAGACAGA 20688
QY 150 TGCTTCAAGTGAACCAAGTTAACTAAG 181
DB 20687 GTTTAAAGATGACCAACCAATTAGATGAG 20656

RESULT 3
US-09-710-279-1929
; Sequence 1929, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1929
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1929

Query Match 10.8%; Score 36.2; DB 4; Length 1587;
Best Local Similarity 53.1%; Pred. No. 0.54;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 141 CAAACGTGCTTACAGTGAAGAAACAGTTAACTAAAGTACCCCTCTTGACAA 200
DB 203 CACACATGACTTTAAAGTTGAACCTTATTTTACAAAGTACCTCAAGCCCATCTA 262
QY 201 GATCAATGCCAGTTGAGCTTAGCCAGCCACATCATCATGTAATGCTTCTCTGATA 260

DB 263 GAAAAATTAACAAATTATCTCAAGCATATGATCATCATTAATCACTTGAA 322
QY 261 AGCTGTTCAATTAATTCCTTTGCA 285
DB 323 ATAAATACAAATTAATCTTGCA 347

RESULT 4
US-09-134-001C-1667
; Sequence 1667, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1667
; LENGTH: 1878
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1667

Query Match 10.8%; Score 36.2; DB 3; Length 1878;
Best Local Similarity 53.1%; Pred. No. 0.58;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 141 CAAACGTGCTTACAGTGAAGAAACAGTTAACTAAAGTACCCCTCTTGACAA 200
DB 494 CAGAGATGACTTTAAAGTTGAACCTTATTTTACAAAGTACCTCAAGCCCATCTA 553
QY 201 GATCAATGCCAGTGTGAGCTTTACCGCCATCATCATTAATGCTTCTTGATA 260
DB 554 GAAAAATTAACAAATTTCTCAAGCATATGATCATCATTAATCACTTGAAA 613
QY 261 AGCTGTTCAATTAATTCCTTTGCA 285
DB 614 ATAAATACAAATTAATCTTGCA 638

RESULT 5
US-09-710-279-3907
; Sequence 3907, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3907
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3907

Query Match 10.8%; Score 36.2; DB 4; Length 3000;
Best Local Similarity 53.1%; Pred. No. 0.69;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY	14	CAAAAGGTGGCTTAAAGGGAAAACA	GTAACTAAGAGACCCCTCTTGGCAA	207
Db	129	CAAGCATGCTTTAAAGTTGAAACCTT	ATTATTTCAAGAGTTACTCTTACGCCATCTA	188
QY	201	GATCAATGCACAGTGGAGCTTTAGC	ACGCAATCATGTAAATGCTTCTGTATA	267
Db	189	GAAAAATTAACAATAATATCTCAAG	CAATATGATATCATCAATTAACAATTCTGAAA	248
QY	261	AGCCGTTCATAAATTCCTTTTSCA	285	
Db	249	ATTAATATCAACAATTAATTCTTGSCA	273	

```

RESULT 6
US-09-710-279-4169/C
: Sequence 4169, Application US/09710279
: Patent No. 6703492
: GENERAL INFORMATION:
: APPLICANT: KIMMERLY, WILLIAM JOHN
: TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
: FILE REFERENCE: PUS4980US
: CURRENT APPLICATION NUMBER: US/09/710,279
: CURRENT FILING DATE: 2000-11-09
: PRIOR APPLICATION NUMBER: 60/164,258
: PRIOR FILING DATE: 1999-11-09
: NUMBER OF SEQ ID NOS: 4472
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4169
: LENGTH: 3368
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: synthetic
: US-09-710-279-4169

```

	Query Match	10.8%	Score 36.2;	DB 4;	Length 3368;
	Best Local Similarity	53.1%;	Pred. No. 0.72;		
	Matches 77;	Conservative 0;	Mismatches 68;	Indels 0;	Gaps 0;
QY	141 CAAACAGTGTGCTTACAGTGGAAAAACAAGTTAAACTAAGTAGACCCCTCTTGACAA				200
Db	1038 CAGACGATGACTTTTAAAGTTGAAAACTTATTTTTCAGAGATTACTTCAAGCCCATACTA				979
QY	201 GATCAATGCCACAGTTGAGCTTTAGCCGACACATCATCATGTAATTCCTTCCGTATA				260
Db	978 GAAAAATPAAACAATTATCTCAAGCATATGATGATATCATCATATATCAACTTGTAAA				919
QY	261 AGCTGTTCAATTAATTCCTTTGCA				285
Db	918 ATAAATATCAACAATTACTTGTGCA				894

```

RESULT 7
US-09-949-016-13002/C
; Sequence 13002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

? SEQ ID NO 13002
? LENGTH: 168394
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)....(168394)
? OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13002
```

Query Match	10.7%;	Score 36;	DB 4;	Length 168394;
Best Local Similarity	58.3%;	Pred. No. 3.7;		
Matches 63;	Conservative 0;	Mismatches 45;	Indels 0;	Gaps 0;

QY 27 TTTTATCATCAACCTCAATTAAGATCAACAGATGATGGGTAGTGACTGTTATTAAGA 86
Db 124517 TTTTGACCTCAATTAAGAGTAAAGATCAATAGAAATAAGAAAAAATAGTTGCTCTTTATTAAGA 124458
QY 87 AGAATAATTAAGATCACTATCATCACTTTAGAGCAATTAAGAGGAGGAG 134
Db 124457 TTTATTTATTTATTAATTTCTCTCAACATAGTATTAAGGAGATCAAG 124410

RESULT 8
US-08-232-463-14/c
; Sequence 14, Application US/08232463

GENERAL INFORMATION:
APPLICANT: DORNER, P.
APPLICANT: SCHEIFLINER, F.
APPLICANT: PALMER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria

```

: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patent in Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/232,463
:
: FILING DATE:
:

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,113
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:

```

1  REGISTRATION NUMBER: 29, 768
2  REFERENCE/DOCKET NUMBER: 30472/114 IMMU
3  TELECOMMUNICATION INFORMATION:
4  TELEPHONE: (703) 836-9300
5  TELEFAX: (703) 683-4109
6  TELEX: 899149
7  INFORMATION FOR SEQ ID NO. 14:
8  SEQUENCE CHARACTERISTICS:
9  LENGTH: 7218 base pairs
10  TYPE: nucleic acid
11  STRANDEDNESS: single
12  TOPOLOGY: linear
13  IMMEDIATE SOURCE:
14  CLONE: pTZ19c-F15
15  OS-08-232-463-14

```

Query Match	10.7%; Score 35.8; DB 1; Length 7218;
-------------	---------------------------------------

```
Best Local Similarity 7.4%; Pred. No. 1.3;
Matches 13; Conservative 100; Mismatches 62; Indels 0; Gaps 0;

QY 10 TTCATTAGATAGCCGTTTTCATCAACTCAATAAGATGAACGAATGATGGTTA 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1459 TTTAAAGATAGAAATTTGGTACRCRRRRRRRRRRRRRRRRRRRRRRRR 1400
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 GTGACGTTTATTAAGAAGTAAATAAGTACTATCATCTTTGAGCAATGAGGAG 129
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1399 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1340
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 GAGAGTTGACGAACAGTGTCTTCAAGTGAATAAAGTTAACTAAAGTGA 184
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1339 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1285
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-09-949-016-44395/C
; Sequence 44395, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44395
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44395

Query Match 10.6%; Score 35.6; DB 4; Length 601;
Best Local Similarity 57.4%; Pred. No. 0.57;
Matches 62; Conservative 1; Mismatches 45; Indels 0; Gaps 0;

QY 27 TTTTCTACACTCAATAAGATGAACGAATGATGGTTGCTACTGTTTATAAGA 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 TTTTGACCTCAATAAGATGAATCAATGAATAAGAAATGCTCTTTATAGAGT 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 AGAGTATAAAGATCATCATCATTTGAGCAATAAGGAGGAGAG 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 TTATTATATTATTAATTCCTTAACAATGAGTAGTAAGAGTCAGAG 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-09-949-016-58898
; Sequence 58898, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 58898
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-58898

Query Match 10.6%; Score 35.6; DB 4; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.57;
Matches 77; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 43 AATAAGATGAACAGATGAATGGTTAGTGACTGTTTATAAGAAGTAAATGAATAC 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38132 ATTTCATGAGATGCGAGATGATGCTTAATGCTTTTGAATATTTATACCAATGAAGAT 38191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 TATCATCTTTGAGCAATTAAGGAGGAGAGATTCAGCAACAGTGTGCTTACAGTGG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38192 TATATGATGCCAGAGATTAATGACCAAGAGCAGAGAAAGAGAGGACAGCAAGTTC 38251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 AAAACAGTTAAACTAAAGTGACCCC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38252 AAAAGCTGTAGAGGCCAGAGAGACC 38277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-949-016-12822/C
; Sequence 12822, Application US/09949016
; Patent No. 6812339
```

Query Match	Best Local Similarity	Score	DB	Length
10.64	49.04	35.6	4	256171
Matches	Conservative	0	Mismatches	99
95	0	Indels	0	Gaps
				0

QY	4	ATATATTTCAATGATAGCGGTTTTTTTCTCAACTCAAAATGAAGTGAACGAATGAT	63
Db	88458	ACTGATTTTATCATTTACACACTTACATTCATTAAGAAATATCACTGATATCTCATTAAT	883999
QY	64	GGGTTTAGTACTGGTTTTATAAGAGAATATATAAGATATCATCATTTTGAAGCAATTA	123
Db	88398	ATGTAATAATATGTGATCAATTTTAAAAAACCAAAAAATAATTTTCTTATGATCATGTA	88339
QY	124	GGGAGGGAGAGATTACGCAACAGTGTGCTTCAAGTGAATAAACAATTAACTTAAAGT	183
Db	88338	ATTAATACATACCCATGTTAAAAAGTTGGAAATTCACAGAAACATGATGAAAAAATTA	88279
QY	184	ACCCCTCTCTTGA	197
Db	88278	AAACTACCATTTAA	88265

Query Match	10.5%;	Score 35.2;	DB 4;	Length 601;
Best local Similarity	48.5%;	Pred. No. 0.75;		
Matches	94;	Conservative	1;	Mismatches 99; Indels 0; Gaps 0;
QY	4	ATTATTTTCATTAGATAGCCGGTTTTTTTCTACAACTCAATAAGATGAACAGATGAAT	63	
Db	232	ACTGATTTTCATCATTTACACACCTTACATCATTAAGAAAATATCAGATGATTCATTAAT	291	
QY	64	GGGTTAGTGACTGTTTATTAAGAAGAGTAATAAGATCTATCATCATTTTGAGCAATAA	123	
Db	292	ANGTAAAAATRTTGATGATCAATTTTAAAAACCAAAAACTATTTTCTTAGATCATGTAA	351	
QY	124	GGGAGGAGAAAGATTACGCAAAACAGTGTCTTCAAGTGGAAAAACAGTTAACTAAAGTG	183	
Db	352	ATAATAACATACCCAGTGTAAAAAGTTGGAATTCACAGAAAGAACATGATGATAAAAAATT	411	
QY	184	ACCCCCCTCCTTGA	197	
Db	412	AAACTRACCATTTAA	425	

RESULT 15
US-08-998-416-187/c
; Sequence 187, Application US/08998416
; Patent No. 6239264

GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-Dec-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-Dec-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGCI976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 663 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1074UP
US-08-998-416-187

Query Match 10.5%; Score 35.2; DB 3; Length 663;
Best local Similarity 58.7%; Pred. No. 0.78; Mismatches 43; Indels 0; Gaps 0;
Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 1 ATTATTATTCTTCAATAGATGCGGTTTCTACTCAACTCAATAGATGAACAGATG 60
DB 567 ATTATTAACCTTATAGTTAAACCATTTATTAATGATCATTAATATAATAAGAAATA 508
QY 61 AATGGTTAGTGTCTGTTTATAAGAGACTATAAGATCTA 104
DB 507 CATTAAATGATATATAGTTATTAAGAACCAATGAAGATACTA 464

Search completed: November 21, 2005, 08:29:58
Job time : 86.8113 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 07:00:52 ; Search time 537.283 Seconds
(without alignments)
5171.686 Million cell updates/sec

Title: US-10-099-663-2
Perfect score: 336
Sequence: 1 attattatttcttagata.....agacagaagatgcattta 336

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9794790 segs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
18:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
19:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
20:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
21:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
22:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
23:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
24:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
25:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
26:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
27:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
28:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	100.0	336	US-10-099-663-2	Sequence 2, Appli
2	336	100.0	2381	US-10-099-663-1	Sequence 1, Appli
3	42.8	12.7	558	US-09-925-065A-572321	Sequence 572321,
4	42.8	12.7	679	US-09-925-065A-875384	Sequence 875384,
5	42.8	12.7	679	US-09-925-065A-875385	Sequence 875385,

C	6	42.8	12.7	679	13	US-09-925-065A-875386	Sequence 875386,
	7	40.4	12.0	116327	21	US-10-719-993-6867	Sequence 6867, Ap
	8	39	11.6	2933	26	US-11-097-143-25165	Sequence 25165, A
	9	37.4	11.1	593	13	US-09-925-065A-115086	Sequence 115086,
	10	37.4	11.1	593	13	US-09-925-065A-115087	Sequence 115087,
	11	37.4	11.1	593	13	US-09-925-065A-115089	Sequence 115089,
	12	37.2	11.1	2927	26	US-11-097-143-7087	Sequence 7087, Ap
	13	37.2	11.1	123526	10	US-09-910-185-11	Sequence 11, Appl
	14	37	11.0	593	13	US-09-925-065A-115088	Sequence 115088,
	15	37	11.0	1846	13	US-09-925-065A-701240	Sequence 701240,
	16	36.8	11.0	98439	22	US-10-741-600-17724	Sequence 17724, A
	17	36.6	10.9	403	9	US-09-815-343-915	Sequence 915, App
	18	36.6	10.9	403	19	US-10-097-105-915	Sequence 915, App
	19	36.6	10.9	3322	26	US-11-097-143-13510	Sequence 13510, A
	20	36.2	10.8	621	13	US-09-925-065A-208033	Sequence 208033,
	21	36.2	10.8	1878	24	US-10-724-972A-1270	Sequence 1270, Ap
	22	36.2	10.8	8648	18	US-10-221-613-217	Sequence 217, App
	23	35.8	10.7	597	13	US-09-925-065A-338035	Sequence 338035,
	24	35.8	10.7	714	22	US-10-706-635-64	Sequence 64, Appl
	25	35.8	10.7	50000	22	US-10-706-635-26	Sequence 26, Appl
	26	35.6	10.6	598	13	US-09-925-065A-350985	Sequence 350985,
	27	35.6	10.6	165097	24	US-10-737-082-77	Sequence 77, Appl
	28	35.6	10.6	165097	24	US-10-765-790-77	Sequence 77, Appl
	29	35.4	10.5	528	13	US-09-925-065A-830377	Sequence 830377,
	30	35.4	10.5	367378	17	US-10-312-841-1	Sequence 1, Appli
	31	35.2	10.5	479	18	US-10-242-535A-47091	Sequence 47091, A
	32	35.2	10.5	479	19	US-10-085-783A-47091	Sequence 47091, A
	33	35.2	10.5	1959	13	US-09-925-065A-28795	Sequence 28795, A
	34	35.2	10.5	1959	13	US-09-925-065A-28796	Sequence 28796, A
	35	35.2	10.5	1959	13	US-09-925-065A-28797	Sequence 28797, A
	36	35.2	10.5	8622	16	US-10-311-455-2115	Sequence 2115, Ap
	37	35.2	10.5	15714	16	US-10-311-455-1146	Sequence 1146, Ap
	38	35.2	10.5	15714	20	US-10-433-793-88	Sequence 88, Appl
	39	35	10.4	134738	14	US-10-087-192-502	Sequence 502, App
	40	34.8	10.4	478	13	US-09-925-065A-178357	Sequence 178357,
	41	34.8	10.4	478	13	US-09-925-065A-178358	Sequence 178358,
	42	34.8	10.4	603	13	US-09-925-065A-345133	Sequence 345133,
	43	34.8	10.4	647	13	US-09-925-065A-345139	Sequence 345139,
	44	34.8	10.4	75899	9	US-09-854-883-243	Sequence 243, App
	45	34.8	10.4	75899	18	US-10-360-510-243	Sequence 243, App

ALIGNMENTS

RESULT 1
US-10-099-663-2
; Sequence 2, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian Gut-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(336)
; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region
US-10-099-663-2

Query Match 100.0%; Score 336; DB 17; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.7e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATTATTATTCTTAGATGACCGCTTTTCTCACTCACTCAATGATGACATGATG 60
|||||

```
Db      1 ATTATATTTCATTAGATAGCCGGTTTTTACTACAACTCAATTAAGATGACAGATG 60
Qy      61 AATGGTTGTACCTTTTATAAGAGATTAAGTACTATCATATTGGAGCA 120
Db      61 AATGGTTGTACCTTTTATAAGAGATTAAGTACTATCATATTGGAGCA 120
Qy      121 TAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGGAACAAAGTTAACTAA 180
Db      121 TAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGGAACAAAGTTAACTAA 180
Qy      181 GTGACCCCTCTCTTGACAAAGATCAATGCAAGTTAGCTTACGACCATCATCA 240
Db      181 GTGACCCCTCTCTTGACAAAGATCAATGCAAGTTAGCTTACGACCATCATCA 240
Qy      241 TGTAAATGCTTCTCGATTAAGCTGTTCATAAATCTCTTTGCAAAAGCTGTGACTTA 300
Db      241 TGTAAATGCTTCTCGATTAAGCTGTTCATAAATCTCTTTGCAAAAGCTGTGACTTA 300
Qy      301 CCAGAACTGTGCTTACAGACAGAAAGATGACATTTA 336
Db      301 CCAGAACTGTGCTTACAGACAGAAAGATGACATTTA 336

RESULT 2
US-10-099-663-1
; Sequence 1, Application US/10099663
; Publication No. US2003017516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian GUT-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; LENGTH: 2381
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(1626)
; NAME/KEY: exon
; LOCATION: (1627)..(1693)
; NAME/KEY: Intron
; LOCATION: (1694)..(2322)
; NAME/KEY: exon
; LOCATION: (2323)..(2381)
; US-10-099-663-1

Query Match      100.0%; Score 336; DB 17; Length 2381;
Best Local Similarity 100.0%; Pred. No. 4,1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATTATATTTCATTAGATAGCCGGTTTTTACTACAACTCAATTAAGATGACAGATG 60
Db      1301 ATTATATTTCATTAGATAGCCGGTTTTTACTACAACTCAATTAAGATGACAGATG 1360
Qy      61 AATGGTTGTACCTTTTATAAGAGATTAAGTACTATCATATTGGAGCA 120
Db      1361 AATGGTTGTACCTTTTATAAGAGATTAAGTACTATCATATTGGAGCA 1420
Qy      121 TAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGGAACAAAGTTAACTAA 180
Db      1421 TAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGGAACAAAGTTAACTAA 1480
Qy      181 GTGACCCCTCTCTTGACAAAGATCAATGCAAGTTAGCTTACGACCATCATCA 240
Db      1481 GTGACCCCTCTCTTGACAAAGATCAATGCAAGTTAGCTTACGACCATCATCA 1540
Qy      241 TGTAAATGCTTCTCGATTAAGCTGTTCATAAATCTCTTTGCAAAAGCTGTGACTTA 300
Db      1541 TGTAAATGCTTCTCGATTAAGCTGTTCATAAATCTCTTTGCAAAAGCTGTGACTTA 1600
```

```
Qy      301 CCAGAACTGTGCTTACAGACAGAAAGATGACATTTA 336
Db      1601 CCAGAACTGTGCTTACAGACAGAAAGATGACATTTA 1636

RESULT 3
US-09-925-065A-572321/C
; Sequence 572321, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 572321
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-572321

Query Match      12.7%; Score 42.8; DB 13; Length 558;
Best Local Similarity 53.6%; Pred. No. 0.11;
Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy      1 ATTATATTTCATTAGATAGCCGGTTTTTACTACAACTCAATTAAGATGACAGATG 60
Db      346 ATTAACTTTAACTAATGATGACCTTTTATATCTCTTTAAACAGAACTTTAGAACT 287
Qy      61 AATGGTTGTACCTTTTATAAGAGATTAAGTACTATCATCATTTGGAGCA 120
Db      286 TAAGAGTAAGTCATGATTAATTAAGATTAAGAACTTATTAAGATTAAGCA 227
Qy      121 TAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGGAACAA 166
Db      226 TAAGAACTAATAATTAAGAAATAATTAATAAAATGTACAA 181

RESULT 4
US-09-925-065A-875384/C
; Sequence 875384, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 875384
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-875384

Query Match      100.0%; Score 336; DB 17; Length 2381;
Best Local Similarity 100.0%; Pred. No. 4,1e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATTATATTTCATTAGATAGCCGGTTTTTACTACAACTCAATTAAGATGACAGATG 60
Db      1301 ATTATATTTCATTAGATAGCCGGTTTTTACTACAACTCAATTAAGATGACAGATG 1360
Qy      61 AATGGTTGTACCTTTTATAAGAGATTAAGTACTATCATATTGGAGCA 120
Db      1361 AATGGTTGTACCTTTTATAAGAGATTAAGTACTATCATATTGGAGCA 1420
Qy      121 TAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGGAACAAAGTTAACTAA 180
Db      1421 TAAGGAGGAGAGATTTCAGCAACAGTGTCTTACAACTGGAACAAAGTTAACTAA 1480
Qy      181 GTGACCCCTCTCTTGACAAAGATCAATGCAAGTTAGCTTACGACCATCATCA 240
Db      1481 GTGACCCCTCTCTTGACAAAGATCAATGCAAGTTAGCTTACGACCATCATCA 1540
Qy      241 TGTAAATGCTTCTCGATTAAGCTGTTCATAAATCTCTTTGCAAAAGCTGTGACTTA 300
Db      1541 TGTAAATGCTTCTCGATTAAGCTGTTCATAAATCTCTTTGCAAAAGCTGTGACTTA 1600
```

NUMBER OF SEQ ID NOS: 957086
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 875384
 LENGTH: 679
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-925-065A-875384

Query Match
 Best Local Similarity 12.7%; Score 42.8; DB 13; Length 679;
 Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

1 ATTATATTTTCTAGATAGCGGTTTCTTACCACTCAATATAGATGAGCAATG 60
 515 ATTAACTTAAACAATTAGAGCTTTATATCTCTTAAACAAGCACTTTAGAGACT 456
 61 AATGGTTAGTGAAGCTTTTAAAGAGAGTAATTAAGATCATCATCTTTAGAGCA 120
 455 TAAGAGTAAGTCCATGTATATTAAGATTAAGAACTATATTAAGATTAAGCA 396
 121 TAAGGAGGAGAGGATTCAGCAACAGTGTCTTACAGTGAATA 166
 395 TAAGAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 350

RESULT 5

US-09-925-065A-875385/c
 Sequence 875385, Application US/09925065A
 Publication No. US20050228172A9
 GENERAL INFORMATION:

APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single
 FILE REFERENCE: 108627.135
 CURRENT APPLICATION NUMBER: US/09/925,065A
 CURRENT FILING DATE: 2001-08-08
 PRIOR APPLICATION NUMBER: US 60/243,096
 PRIOR FILING DATE: 2000-10-24
 PRIOR APPLICATION NUMBER: US 60/252,147
 PRIOR FILING DATE: 2000-11-20
 PRIOR APPLICATION NUMBER: US 60/250,092
 PRIOR FILING DATE: 2000-11-30
 PRIOR APPLICATION NUMBER: US 60/261,766
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/289,846
 PRIOR FILING DATE: 2001-05-09
 NUMBER OF SEQ ID NOS: 957086
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 875385
 LENGTH: 679
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-925-065A-875385

Query Match
 Best Local Similarity 12.7%; Score 42.8; DB 13; Length 679;
 Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

1 ATTATATTTTCTAGATAGCGGTTTCTTACCACTCAATATAGATGAGCAATG 60
 515 ATTAACTTAAACAATTAGAGCTTTATATCTCTTAAACAAGCACTTTAGAGACT 456
 61 AATGGTTAGTGAAGCTTTTAAAGAGAGTAATTAAGATCATCATCTTTAGAGCA 120
 455 TAAGAGTAAGTCCATGTATATTAAGATTAAGAACTATATTAAGATTAAGCA 396
 121 TAAGGAGGAGAGGATTCAGCAACAGTGTCTTACAGTGAATA 166
 395 TAAGAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 350

RESULT 6
 US-09-925-065A-875386/c

Sequence 875386, Application US/09925065A
 Publication No. US20050228172A9
 GENERAL INFORMATION:

APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single
 FILE REFERENCE: 108627.135
 CURRENT APPLICATION NUMBER: US/09/925,065A
 CURRENT FILING DATE: 2001-08-08
 PRIOR APPLICATION NUMBER: US 60/243,096
 PRIOR FILING DATE: 2000-10-24
 PRIOR APPLICATION NUMBER: US 60/252,147
 PRIOR FILING DATE: 2000-11-20
 PRIOR APPLICATION NUMBER: US 60/250,092
 PRIOR FILING DATE: 2000-11-30
 PRIOR APPLICATION NUMBER: US 60/261,766
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/289,846
 PRIOR FILING DATE: 2001-05-09
 NUMBER OF SEQ ID NOS: 957086
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 875386
 LENGTH: 679
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-925-065A-875386

Query Match
 Best Local Similarity 12.7%; Score 42.8; DB 13; Length 679;
 Matches 89; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

1 ATTATATTTTCTAGATAGCGGTTTCTTACCACTCAATATAGATGAGCAATG 60
 515 ATTAACTTAAACAATTAGAGCTTTATATCTCTTAAACAAGCACTTTAGAGACT 456
 61 AATGGTTAGTGAAGCTTTTAAAGAGAGTAATTAAGATCATCATCTTTAGAGCA 120
 455 TAAGAGTAAGTCCATGTATATTAAGATTAAGAACTATATTAAGATTAAGCA 396
 121 TAAGGAGGAGAGGATTCAGCAACAGTGTCTTACAGTGAATA 166
 395 TAAGAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 350

RESULT 7

US-10-719-993-6867
 Sequence 6867, Application US/10719993
 Publication No. US20040265849A1
 GENERAL INFORMATION:

APPLICANT: Cargill, Michele et al.
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 FILE REFERENCE: CL001496
 CURRENT APPLICATION NUMBER: US/10/719,993
 CURRENT FILING DATE: 2003-11-24
 NUMBER OF SEQ ID NOS: 55342
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 6867
 LENGTH: 116327
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (1)...(116327)
 OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
 US-10-719-993-6867

Query Match
 Best Local Similarity 12.0%; Score 40.4; DB 21; Length 116327;
 Matches 77; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

197 ACAAGATCAATGCCAGTGTAGCTTTAGCCGACCAATCATCATTAATGCTTCT 256

```
Db 70844 AAAAGATGATGCCCTCACTAGGTTTGACACAGTATTTAGCATGATTAAGATTACCT 70903
Qy 257 GATTAAGCCGTTCATTAATTCCTTTGCAAGCTGCTACTACGAGAAGTCTGCTAC 316
Db 70904 GGTAACTTTTAAATAATTCATTTCCCAAGTCTTAATTAATAAAGAGTCAGTTT 70963
Qy 317 AGACAGAAAGATGTCAT 334
Db 70964 AGAAGAAAAAGATATT 70981

RESULT 8
US-11-097-143-25165
; Sequence 25165, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25165
; LENGTH: 2933
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-25165

Query Match 11.6%; Score 39; DB 26; Length 2933;
Best Local Similarity 52.1%; Pred. No. 3;
Matches 87; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 73 ACTGTTTAAAGACAGTAATTAAGATCTATCATCTTTGAGCAATTAAGGAGGAG 132
Db 2766 AATTTAAAGACAGATCTTACTTAAGAAATTAAGATTAATGATGAGAAAGAGAGAT 2825
Qy 133 AGATTGACAAACAGTGTGCTTACAGTGGAAAAACAAGTTAACTAAGTACCCTCC 192
Db 2826 AATTTAAAGCAAAATTTCTTATGATGGAACAGAGAACCCGAAATTTGCATTATC 2885
Qy 193 CTGACAAAGTCAATCCACAGTTGAGCTTTAGCCAGCCACATCATC 239
Db 2886 CGTGGCCCACTGAATCATCTTATGTGACACCAACCATACCAACC 2932

RESULT 9
US-09-925-065A-115086
; Sequence 115086, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
```

```
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115086
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-115086

Query Match 11.1%; Score 37.4; DB 13; Length 593;
Best Local Similarity 49.8%; Pred. No. 4.1;
Matches 123; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

Qy 52 AACAGAAATGAATGGTTAGTGCCTTTTAAAGAAAGTATTAAGATATCATCAT 111
Db 290 AACAGAAATGAATATTTTAAAGATATTAAGTATTAAGAAAGATTAAGATATCAT 349
Qy 112 TTGAGGCAATTAAGGAGGAGGAGATTGACCAACAGTGTCTTACAGTGAACAACT 171
Db 350 TTTTCAATATGATATAGCAATTTAATCTGAACAAACAGAGACTGAATGAATTTATT 409
Qy 172 TTAACCTAAAGTACCCCTCTTGAACAAGATCAATGCCACGTTGAGCTTGAAGCC 231
Db 410 AAACCTGATTTTAAAGTCACTAAATGACCAATGCCAAATGCCAA--GACACATGAATCA 466
Qy 223 AATCATCATGTAATTTGCTTCCTGATTAAGCTGTTCATTAATCTCTTGCAGAACTC 291
Db 467 AAGCTTTTGGACACTAATTAATTAAGGACATTTTCTAATTTAATTTCTTTACAAAGTA 526
Qy 292 TGCTACT 298
Db 527 AGCAGCT 533

RESULT 10
US-09-925-065A-115087
; Sequence 115087, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115087
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Homo sapiens
```


US-09-925-065A-115087

Query Match 11.1%; Score 37.4; DB 13; Length 593;
Best Local Similarity 49.8%; Pred. No. 4.1;
Matches 123; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

QY 52 AACAGATGATGAGTGGTGTAGTCTTTATTAAGAGATTAATTAAGTACTTCATCAT 111
DB 290 AACAGAAAAAGATATTTTAAAGATTAATTAAGAGAGAGATTAATTAATCATCAT 349
QY 112 TTGAGCAATTAAGGAGGAGAGATTCAGCAAAAGTGTGCTTACAAAGTGAAGAAAGT 171
DB 350 TTTTGCAATGATATGATGCAATTTTAAGTAAAGACAAAGAGCTTAATGAAATTTAT 409
QY 172 TAACTTAAGTGAACCCCTCTCTTGAAGATCAATGCCAGCTTGAAGCTTGAAGCC 231
DB 410 AAAACGATTTTAAAGTTCAGTAAATGCGACAAGGCCAAA--GACACATAGGAATCA 466
QY 232 ACATCATCATGTAATTTGCTTCTGTAAGCCTGTTCAATTAATTCCTTTGCAAGCTC 291
DB 467 ACAGCTTTTTCGACACTAAATTAAGGAGATTTCTAAATTTAAATTTCTTTACAAAGTA 526
QY 292 TGCTACT 298
DB 527 AGCAGCT 533

RESULT 11

US-09-925-065A-115089

; Sequence 115089, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115089
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-115089

Query Match 11.1%; Score 37.4; DB 13; Length 593;
Best Local Similarity 49.8%; Pred. No. 4.1;
Matches 123; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

QY 52 AACAGATGATGAGTGGTGTAGTCTTTATTAAGAGATTAATTAAGTACTTCATCAT 111
DB 290 AACAGAAAAAGATATTTTAAAGATTAATTAAGAGAGAGATTAATTAATCATCAT 349
QY 112 TTGAGCAATTAAGGAGGAGAGATTCAGCAAAAGTGTGCTTACAAAGTGAAGAAAGT 171
DB 350 TTTTGCAATGATATGATGCAATTTTAAGTAAAGACAAAGAGCTTAATGAAATTTAT 409
QY 172 TAACTTAAGTGAACCCCTCTCTTGAAGATCAATGCCAGCTTGAAGCTTGAAGCC 231
DB 410 AAAACGATTTTAAAGTTCAGTAAATGCGACAAGGCCAAA--GACACATAGGAATCA 466
QY 232 ACATCATCATGTAATTTGCTTCTGTAAGCCTGTTCAATTAATTCCTTTGCAAGCTC 291

DB 467 ACAGCTTTTTCGACACTAAATTAAGGAGATTTCTAATTAATTAATTCCTTTACAAAGTA 526

QY 292 TGCTACT 298
DB 527 AGCAGCT 533

RESULT 12

US-11-097-143-7087

; Sequence 7087, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; DROSOPHILA GENES.
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7087
; LENGTH: 2927
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-7087

Query Match 11.1%; Score 37.2; DB 26; Length 2927;
Best Local Similarity 51.9%; Pred. No. 9.7;
Matches 84; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 73 ACTGTTTATTAAGAGATTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 132
DB 2766 ACATTAATGACAGATCTTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 2825
QY 133 AGATTCAGCAACAGAGTGTCTTACAGTGAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 192
DB 2826 ATATTAACAGAAATTTCTTAATGATGTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2885
QY 193 CTTGACAAATCAATGCCAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 234
DB 2886 CGTGCCCACTGAATGCAATCACTTAATGAGACACCAACCAATA 2927

RESULT 13

US-09-910-185-11/C

; Sequence 11, Application US/09910185
; Publication No. US20030083279A1
; GENERAL INFORMATION:

; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freiler
; TITLE OF INVENTION: ANTISENSE MODULATION OF GLIOMA-ASSOCIATED ONCOGENE-3 EXPRESSION
; FILE REFERENCE: RTS-0258
; CURRENT APPLICATION NUMBER: US/09/910,185
; PRIOR FILING DATE: 2001-07-18

```
/ NUMBER OF SEQ ID NOS: 90
/ SEQ ID NO 11
/ LENGTH: 123526
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-910-185-11
```

```
Query Match      11.1%; Score 37.2; DB 10; Length 123526;
Best Local Similarity 53.4%; Pred. No. 53;
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
```

```
QY      12 CATTAGATGACCGGTTTCTTACTACACTCAATTAAGTGAACAGAAATGGGTTACT 71
      107264 CATTCCCAATGCTTTTAAATGCTACAGAAATGCAAGAGAGGGGTGCT 107205
DB      72 GACTGTTTAAAGAGAGAAATTAAGTACTCTCATTTGAGGCATTAAGGAGGGA 131
      107204 GTGTACCCCTTAAATTAATGAGATGAAGAAAGAAAGGCAAGAGGAGGA 107145
QY      132 GAGATTCAGCAACAGTGTGCTTACA 157
      107144 GGGAGAAACCCGAAAGCATTCATGCA 107119
DB
```

```
RESULT 14
US-09-925-065A-115088
/ Sequence 115088, Application US/09925065A
/ Publication No. US2005028172A9
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 115088
/ LENGTH: 593
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-115088
```

```
Query Match      11.0%; Score 37; DB 13; Length 593;
Best Local Similarity 49.4%; Pred. No. 5.3;
Matches 122; Conservative 1; Mismatches 121; Indels 3; Gaps 1;
```

```
QY      52 AACGAGATGATGGCTTATGACTGTTTAAAGAGAGTAATTAAGTACTATCAT 111
      290 AACAGAAAAAGATATTTTAAAGATTAATCTAGAAAAGAGATTAATCATCAT 349
DB      112 TTGAGCAATAGAGGAGGAGATTCAGCAACAGTGTCTTACAAGTGAACAAG 171
      350 TTTTGCAATGATATAGCAATTTTACTAGAAAACACAGAGACTGAATGAAATTTAT 409
QY      172 TAACTAAAGTGAACCCCTCTCTGACAGATCATATGCCACAGTTGAGCTTAC 231
      410 AAAACTGATTTTAAAGTGAATGAATGCAACATGCCAAA---GACACAYAGGAATCA 466
DB      232 ACATCATCATGTAATTTGCTTCTGATTAAGCTGTTCATTAATTTCTTTGAAAGCT 291
      467 ACAGCTTTTTCAGACTAAATTAAGCGAGATTTCTAATTTAAATTTCTTTACAAAAGTA 526
```

```
QY      292 TGCTACT 298
      527 AGCACT 533
DB
```

```
RESULT 15
US-09-925-065A-701240/c
/ Sequence 701240, Application US/09925065A
/ Publication No. US2005028172A9
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 701240
/ LENGTH: 1846
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-701240
```

```
Query Match      11.0%; Score 37; DB 13; Length 1846;
Best Local Similarity 52.2%; Pred. No. 8.9;
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```
QY      16 AGATAGCCCGTTTTTCTACCACTCAATTAAGTGAACAGAAATGGGTTAGTACT 75
      1271 AGATACCTTCATCTTAATTAATCTCTAACCATTAAAGTGATCAAGGAAATTAAGT 1212
DB      76 GTTATTAAGAGAGATTAATTAAGATCTATCATTTTGAGGCAATTAAGGAGGAGAGA 135
      1211 GAATTAACATAAAAACTTATGATTCATACATTTTTCAGAAATTTCTGATTAAGAGA 1152
QY      136 TTCAGCAACAGTGTGCTTACAGAGTGAACAAGTT 172
      1151 AGTAAATATGCTGAAACCAATAGAAAGAAAT 1115
DB
```

```
Search completed: November 21, 2005, 11:55:02
Job time : 541.283 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2005, 15:48:44 ; Search time 9972 Seconds
(without alignments)
11569.582 Million cell updates/sec

Title: US-10-099-663-1
Perfect score: 2381
Sequence: 1 agctccgcgcgaagaag.....atctgaagctcattctcag 2381

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hhg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	736.8	30.9	2847	5	AY254202 Gallus ga
2	85	923	5	5	AY254202 Gallus ga
3	81	923	5	5	AY254202 Gallus ga
4	78.8	3.3	899	5	AY254202 Gallus ga
5	64.2	2.7	1141	6	AX083744 Sequence
6	62	2.6	3013	10	RATFABP
7	62	2.6	230897	2	AC130496 Rattus no
8	57.8	2.4	1141	6	AX083744 Sequence
9	56.6	2.4	5039	10	MUSFABP
10	51.8	2.2	234081	3	PFMALP2
11	50	2.1	564	10	RATFABP
12	49.8	2.1	619	5	XELIFABP
13	49.8	2.1	5204	6	AX771590 Sequence
14	49.8	2.1	5204	9	HUMFABP
15	49.8	2.1	20000	2	AC008077 Homo sapi
16	49.2	2.1	2000	6	AX655393 Sequence
17	49	2.1	136061	2	AC117821 Mus muscu
18	49	2.1	202804	10	AC021630 Mus muscu
19	49	2.1	249943	3	AE014823 Plasmodiu

20	48.4	2.0	564	10	RATFABP	K01180 Rat intesti
21	48.4	2.0	832	6	AR415176	AR415176 Sequence
22	48.4	2.0	832	6	AX972010	AX972010 Sequence
23	48.4	2.0	832	6	BD110729	BD110729 EST and e
24	48.4	2.0	18876	3	CEP3258	272509 Caenorhabdi
25	48.4	2.0	178045	2	CR762472	CR762472 Danio rer
26	48.4	2.0	197110	9	AC104306	AC104306 Homo sapi
27	48.2	2.0	132254	3	AC116330	AC116330 Dictyoste
28	48.2	2.0	132254	3	AC116330	AC116330 Dictyoste
29	48	2.0	191840	9	AC092656	AC092656 Homo sapi
30	47.8	2.0	110000	3	AC116984	Continuation (2 of
31	47.8	2.0	137174	2	BX276186	Continuation (2 of
32	47.8	2.0	162810	2	CR392028	Continuation (2 of
33	47.6	2.0	53302	3	AC116961	Continuation (2 of
34	47.6	2.0	143364	2	BX324220	Continuation (2 of
35	47.6	2.0	174176	5	AL772340	Continuation (2 of
36	47.6	2.0	177623	2	CR394571	Continuation (2 of
37	47.4	2.0	110000	2	PFMAL13P3	Continuation (10 o
38	47.4	2.0	318221	2	PFMAL13P3	Continuation (10 o
39	47.2	2.0	165797	2	CR847851	Continuation (10 o
40	47	2.0	2636	6	CQ593303	Continuation (10 o
41	47	2.0	81181	2	AC019668	Continuation (10 o
42	47	2.0	176969	3	AC010707	Continuation (10 o
43	47	2.0	181815	5	BX248120	Continuation (10 o
44	47	2.0	184877	5	BX649334	Continuation (10 o
45	47	2.0	192540	3	AC010846	Continuation (10 o

ALIGNMENTS

RESULT 1
AY254202 2847 bp DNA linear VRT 22-APR-2003
LOCUS AY254202
DEFINITION Gallus gallus intestinal fatty acid-binding protein gene, complete cds.

ACCESSION AY254202
VERSION AY254202.1 GI:30060211
KEYWORDS
SOURCE
ORGANISM Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 2847)

REFERENCE
Wang, Q., Li, H., Wang, Y. and Zhao, J.
Cloning and characterization of chicken I-FABP gene
Unpublished
2 (bases 1 to 2847)

AUTHORS
Wang, Q., Li, H., Wang, Y. and Zhao, J.
TITLE
Direct Submision
JOURNAL
Submitted (11-MAR-2003) Animal Science & Technology, Northeast
Agricultural University, 59 Mucai Street, Harbin, Heilongjiang
150030, China

FEATURES

source
1..2847
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
/tissue_type="blood"
/note="Breed: Arbor Acres broiler"
join(1..21..87,720..892,1465..1572,2416..2466)
/product="intestinal fatty acid-binding protein"
join(121..87,720..892,1465..1572,2416..2466)
/note="FABP"
/codon_start=1
/product="intestinal fatty acid-binding protein"
/protein_id="AAPI3101.1"
/db_xref="GI:30060212"
/translation="MAFGTGWKIEKNYERKMEAMGVNVRKLGADNLTATLQGD
GNKFLVKSNNRPTDIEFTGTGVSFEYSGLADGTLSGWNLEGNLTGVGTFRKDKNGKV
LTAVETIGSELICQTVYVEGVAKRIFPKE"
ORIGIN

Query Match 30.9%; Score 736.8; DB 5; Length 2847;
 Best Local Similarity 98.1%; Pred. No. 4e-146;
 Matches 757; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

1613 CTACAGACAGAAAGATGCGATTAAACGCTACTTGGAAAAATAGAGAAATAGAACTATG 1672
 7 CGAGAGACAGAAAGATGCGATTAAACGCTACTTGGAAAAATAGAGAAATAGAACTATG 66

1673 AAAAAATTCATGAGAGCAATGGGTAGCCCTTACTTTTGAATGCTCTTAAAGCAGAT 1732
 67 AAAAAATTCATGAGAGCAATGGGTAGCCCTTACTTTTGAATGCTCTTAAAGCAGAT 126

1733 ACCACATGCGCGGAAATACAACTTAAAGCTTCAATGACATACATGCTGCTAACCTGTC 1792
 127 ACCACTACGCGGAAATACAACTTAAAGCTTCAATGACATACATGCTGCTAACCTGTC 186

1793 TTTGTGTCTGCTATTTTGCCTTGCACATGCGCTGCACTTATTTTGAAGAACTCTAT 1852
 187 TTTGTGTCTGCTATTTTGCCTTGCACATGCGCTGCACTTATTTTGAAGAACTCTAT 246

1853 AAGAGGGGAATTCAGAGAGAGAAAAAATTTGATTTTATTTGCAATGCGATTAATCTTATGC 1912
 247 AAGAGGGGAATTCAGAGAGAGAAAAAATTTGATTTTATTTGCAATGCGATTAATCTTATGC 306

1913 ATTAGCTAATTCAGATAGAGGATTCAGAGAGAAATTTAAATAGAAATTAATGTAAGGA 1972
 307 ATTAGCTAATTCAGATAGAGGATTCAGAGAGAAATTTAAATAGAAATTAATGTAAGGA 366

1973 ATATATTTTGTATAGACTGTTTGAAGAAATTAACAGAGAGGAAATTTGCTGCTTCCACT 2032
 367 ATATATTTTGTATAGACTGTTTGAAGAAATTAACAGAGAGGAAATTTGCTGCTTCCACT 426

2033 TTTTGCAGAACACATGATTTT--GAGTCAATTTTAAATGCTAGTCTTACTTTAAGCTT 2089
 427 TTTTGCAGAACACATGATTTT--GAGTCAATTTTAAATGCTAGTCTTACTTTAAGCTT 486

2090 GTACAAATGCGCTGATATATGATGTAAACATTAATCATATGCTAGTGTAGTGTGA 2149
 487 GTACAAATGCGCTGATATATGATGTAAACATTAATCATATGCTAGTGTAGTGTGA 546

2150 TTACAGGCTGAACACTGCTCAGTGAAGAGTGAAGAGTAAGACTGTAAGTCAAT 2209
 547 TTACAGGCTGAACACTGCTCAGTGAAGAGTGAAGAGTAAGACTGTAAGTCAAT 606

2210 CTGGGCTAAGCTCCCTCACTACAGAAAAAAGTCAATTAATAATGCAACATGATGTTCT 2269
 607 CTGGGCTAAGCTCCCTCACTACAGAAAAAAGTCAATTAATAATGCAACATGATGTTCT 666

2270 ATTTGTTTTCTCTGCTGATGTATATGATATATTTTATTTTATTTTATTTTATTTTATTTT 2329
 667 ATTTGTTTTCTCTGCTGATGTATATGATATATTTTATTTTATTTTATTTTATTTTATTTT 726

2330 TGTGATGAAAAAGAACTTGAAGAGCCACAGATAATCTGAAGCTCACTATTCAAG 2381
 727 TGTGATGAAAAAGAACTTGAAGAGCCACAGATAATCTGAAGCTCACTATTCAAG 778

Db 727 TGTGATGAAAAAGAACTTGAAGAGCCACAGATAATCTGAAGCTCACTATTCAAG 778

RESULT 2
 BX932340 923 bp mRNA linear VRT 30-MAR-2004
 LOCUS BX932340
 DEFINITION Gallus gallus finished cDNA, clone CHEST355022.
 ACCESSION BX932340
 VERSION BX932340.2 GI:46017395
 KEYWORDS
 SOURCE
 ORGANISM
 Gallus gallus (chicken)
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus
 1 (bases 1 to 923)
 Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
 Croning, M.D.R., Davies, R.M., Francis, M.D., Grafham, D.V.,
 Hubbard, S.J., Humphray, S.J., Hunt, P.J., Maddison, M., McLaren, S.R.,
 Niblett, D., Overton, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,
 Tickle, C., and Wilson, S.A.
 Direct Submission
 Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: chickens@bbs.unl.ac.uk
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
 sequencing project.
 This sequence is from the
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
 from a library constructed by Elizabeth Bosch. cDNA was prepared
 from RNA extracted from heart, normalised, and poly A-tailed.
 EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
 pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
 coli DH10B.
 Location/Qualifiers
 1..923
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hixex"
 /db_xref="taxon:9031"
 /clone="CHEST355022"
 /clone_1ib="CSBCHN70"
 /dev_stage="stage 36"

ORIGIN

Query Match 3.6%; Score 85; DB 5; Length 923;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1610 TGCCACAGACAGAAAGATGCGATTAAACGCTACTTGGAAAAATAGAGAAATAGGAAT 1669
 1 TGCCACAGACAGAAAGATGCGATTAAACGCTACTTGGAAAAATAGAGAAATAGGAAT 60

1670 ATGAAAAATTCATGAGAGCAATGGG 1694
 61 ATGAAAAATTCATGAGAGCAATGGG 85

Db 61 ATGAAAAATTCATGAGAGCAATGGG 85

RESULT 3
 BX930518 899 bp mRNA linear VRT 02-FEB-2004
 LOCUS BX930518
 DEFINITION Gallus gallus finished cDNA, clone CHEST640B17.
 ACCESSION BX930518
 VERSION BX930518.1 GI:41631046
 KEYWORDS
 SOURCE
 ORGANISM
 Gallus gallus (chicken)
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus
 1 (bases 1 to 899)
 Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
 Croning, M.D.R., Davies, R.M., Francis, M.D., Grafham, D.V.,
 Hubbard, S.J., Humphray, S.J., Hunt, P.J., Maddison, M., McLaren, S.R.,
 Niblett, D., Overton, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,
 Tickle, C., and Wilson, S.A.
 Direct Submission
 Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: chickens@bbs.unl.ac.uk
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
 sequencing project.
 This sequence is from the
 BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
 from a library constructed by Elizabeth Bosch. cDNA was prepared
 from RNA extracted from small intestine, normalised, and poly A-tailed.
 EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
 pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
 coli DH10B.
 Location/Qualifiers
 1..899
 /organism="Gallus gallus"

FEATURES
 source

ORIGIN

Query Match 3.4%; Score 81; DB 5; Length 899;
Best Local Similarity 100.0%; Pred. No. 9.4e-07;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1614 TACAGACAGAAAGATGCGATTAAACGCTTCTGAGAAAATAGAGAAAATGAGAACTATGA 1673
|||||
DB 11 TACAGACAGAAAGATGCGATTAAACGCTTCTGAGAAAATAGAGAAAATGAGAACTATGA 70
|||||

QY 1674 AAAATTCATGAGCAATGGG 1694
|||||
DB 71 AAAATTCATGAGCAATGGG 91
|||||

RESULT 4
BX933597 896 bp mRNA linear VRT 02-FEB-2004
LOCUS Gallus gallus finished cDNA, clone CHST153f19.
DEFINITION BX933597
ACCESSION BX933597.1 GI:41634125
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Archaeopteryx; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 896)
Boatman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
Cromar, M.D.R., Davies, R.M., Francis, M.D., Graffam, D.V.,
Hubbard, S.J., Humphrey, S.J., Hunt, P.J., Maddison, M., McLaren, S.R.,
Niblett, D., Overton, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,
Tickle, C., and Wilson, S.A.
Direct Submission
Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickes@pm.sanger.ac.uk
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from small intestine, and poly A-trimmed.
EcoRI-NciI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NciI Host: Escherichia
coli DH10B.

FEATURES
source
1..896
/organism="Gallus gallus"
/mol_type="mRNA"
/ecstrain="Compton line 151"
/db_xref="taxon:9031"
/clone="ChST153f19"
/clone_1ib="CSEQCCH56"
/dev_stage="adult"

ORIGIN

Query Match 3.3%; Score 78.8; DB 5; Length 896;
Best Local Similarity 97.6%; Pred. No. 2.8e-06;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1613 CTACAGACAGAAAGATGCGATTAAACGCTTCTGAGAAAATAGAGAAAATGAGAACTATG 1672
|||||
DB 7 CGAGACAGAAAGATGCGATTAAACGCTTCTGAGAAAATAGAGAAAATGAGAACTATG 66
|||||

QY 1673 AAAATTCATGAGCAATGGG 1694
|||||
DB 67 AAAATTCATGAGCAATGGG 88
|||||

RESULT 5
AX083744/c 1141 bp DNA linear PAT 28-FEB-2001
LOCUS Sequence 22 from Patent WO0111061.
DEFINITION AX083744
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1 Kunst, L. and Clemens, S.
Regulation of embryonic transcription in plants
Patent: WO 011061-A 22 15-FEB-2001;
UNIVERSITY OF BRITISH COLUMBIA (CA)
Location/Qualifiers
1..1141
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1..1141
/note="consensus sequence of A.t., L.a., and B.n. FAE1
promoters"

ORIGIN

Query Match 2.7%; Score 64.2; DB 6; Length 1141;
Best Local Similarity 12.7%; Pred. No. 0.0034;
Matches 136; Conservative 379; Mismatches 550; Indels 10; Gaps 3;

QY 830 TTTTCTGCATATATCTCTTCAACATTTAAACCTGGATCTATGAAATCAACAGTTGGG 889
|||||
DB 1129 TTTTCTGCATATATCTCTTCAACATTTAAACCTGGATCTATGAAATCAACAGTTGGG 1070
|||||

QY 890 TAAATTCCTAGACAGACATCACTAGCTAGAGATGAGACAGAAAGAGATTCAC 949
|||||
DB 1069 TAAATTCCTAGACAGACATCACTAGCTAGAGATGAGACAGAAAGAGATTCAC 1010
|||||

QY 950 GAATGGCTATATATATAGAAATACGTAGAGAGTCTCGAATTTAGACTACTATTA 1009
|||||
DB 1009 GAATGGCTATATATATAGAAATACGTAGAGAGTCTCGAATTTAGACTACTATTA 950
|||||

QY 1010 GAGTGAAGACAGATGAGAAATATCATGCAATTTCTGAGCTGAGCACTAGCTGAA 1069
|||||
DB 949 GAGTGAAGACAGATGAGAAATATCATGCAATTTCTGAGCTGAGCACTAGCTGAA 890
|||||

QY 1070 GGTTCCTGAAATGAAACGAG-----TTTCCCAACCTACCTGAGATGTTCAAGTAT 1122
|||||
DB 889 GGTTCCTGAAATGAAACGAG-----TTTCCCAACCTACCTGAGATGTTCAAGTAT 830
|||||

QY 1123 CCTCATCTCATGCTTATATATGAGTAGAATGATTCACCAATTAGATGAGCA 1182
|||||
DB 829 CCTCATCTCATGCTTATATATGAGTAGAATGATTCACCAATTAGATGAGCA 770
|||||

QY 1183 AGCAGAGATTTGTTTATCTGTTGGTAAATACGTTTCTCCAGTTGTAAGACCC 1242
|||||
DB 769 AGCAGAGATTTGTTTATCTGTTGGTAAATACGTTTCTCCAGTTGTAAGACCC 710
|||||

QY 1243 TCCACACAGATATAAGTCTATGCAACAAGAAATGTAATACATCTTCTAGTCTCAT 1302
|||||
DB 709 TCCACACAGATATAAGTCTATGCAACAAGAAATGTAATACATCTTCTAGTCTCAT 650
|||||

QY 1303 TATTAATTTTCAATGATGACCGTTTCTTACTACCACTCAATTAAGATGAGATGAA 1362
|||||
DB 649 TATTAATTTTCAATGATGACCGTTTCTTACTACCACTCAATTAAGATGAGATGAA 590
|||||

QY 1363 TGGGTTAGTACTGTTTATTAAGAGATTAAG--ATACTATCATCTTGGAGCA 1420
|||||
DB 589 TGGGTTAGTACTGTTTATTAAGAGATTAAG--ATACTATCATCTTGGAGCA 530
|||||

QY 1421 TAAAGGAGGAGAGATTCAGCAACAGTGTCTTACAACTGAGAAAACAAGTTAAATA 1480
|||||
DB 1421 TAAAGGAGGAGAGATTCAGCAACAGTGTCTTACAACTGAGAAAACAAGTTAAATA 1480
|||||

```

Db      529  MRMAGTMMWRHNNNNNTDTRYYMMKRMABRTTYYDSMCKNSMWRGNMNRAMKMMWAA 470
QY      1481  GTGACCCCTCTCTGTAACAAGATCATGCGACGTTGACCTTAAAGCAGCATCATCA 1540
Db      469  NNDAGANDHTTWMGNNTMMRMAMNNMAMKRAVCCNNNNNRACVHKKHMMKMTWK 410
QY      1541  TGTAAATGCTCTTCTGATTAAGCCTGTGATTAATCTCTTTGCAAGCTCTGCTACTTA 1600
Db      409  YMKKACNNNNBRAMVAVMMVSRDNTTIDMMMTSLMBHMTVVDYTPMRAMNNNN 350
QY      1601  CCAAGAGTGGCTACAGACAGAAAGATGGCACTTAAAGCTACTTGAAATAAGAGAAA 1660
Db      349  NMNRBCKTTSMMWMDHNMHTCTYGNNTGSAVBMAMSMAAGASNBVTYVMCMMTYM 290
QY      1661  ATGAGAACTATGAAAT--TCATGAAGCAATGGGTAAGCCTTACTTTTGAATGCTT 1719
Db      289  GKTMTNNNNNNKMYRTKTVACNNRYDDTAVMTBKRYKCYAVBYBYBMWKHMM 230
QY      1720  CTAAAGACAGATACACCTACGCGGAATACAACTTAAGCTGTGATGACATCATCT 1779
Db      229  BMRBRABHRSMMNMWVKCNKVMVSMHYHAMRYBKABAVGNNNNKMDMAHHHWCATNN 170
QY      1780  GGTAACTGTCTCTTGTGTCTGCTATTTTGGCTTGACATTTGCCCTGCACTTAATTT 1839
Db      169  NMMWMMYAVMHMHKKGKAAWTNNKTBABDHBAAHYKTYWRYDYCAMCMWNAKAKYR 110
QY      1840  GAAAACCTCTATAGAGGGGAATACAGAGAGAAAACATTCGATTTATTTTGC 1894
Db      109  TANKHMYTDRYVSANTGVRMMMRMCMWYSMNNRYMRKRYTAMMYSM 55

RESULT 6
RATFABP
LOCUS   3013 bp      DNA      linear      ROD 27-APR-1993
DEFINITION
a B2 repetitive element.
ACCESSION
M18080.1 GI:204071
VERSION
B2 repetitive sequence; fatty acid binding protein.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 3013)
Sweetser,D.
JOURNAL
Unpublished (1987)
AUTHORS
Sweetser,D.A., Birkmeier,E.H., Klisak,I.J., Zollman,S.,
Sparkes,R.S., Mohandas,T., Lusis,A.J. and Gordon,J.I.
TITLE
The human and rodent intestinal fatty acid binding protein genes. A
comparative analysis of their structure, expression, and linkage
relationships
JOURNAL
J. Biol. Chem. 262 (33), 16060-16071 (1987)
MEDLINE
80858967
PUBMED
2824476
COMMENT
Original source text: Rat (Sprague Dawley) intestinal epithelial
lining DNA, clone lambda-RIFABP.
Draft entry and computer-readable sequence for [2] kindly provided
by D. Sweetser, 19-JAN-1988.
FEATURES
source
1..3013
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
repeat_region
195..406
/feature="B2 repeat"
prim_transcript
1155..>3013
/note="FABPI mRNA (alt.) and introns"
prim_transcript
1179..>3013
/note="FABPI mRNA (alt.) and introns"
join(1215..1281,2541..2713)
/note="fatty acid binding protein"

```

```

/codon_start=1
/protein_id="AA041133.1"
/db_xref="GI:204072"
/translation="MADGTWKVDRNENYKEMKMGINVVYRKUGAHNDNLKLTITOE
GNKFTVKSNSFRNIDVFEIGVDPAVSLAGTELT"
<1215..1281
/note="fatty acid binding protein"
/number=1
1282..2540
/note="FABPI intron A"
2541..2713
/note="fatty acid binding protein"
/number=2
2714..>3013
/note="FABPI intron B"
ORIGIN
1 bp upstream of EcoRI site.
Query Match 2.6%; Score 62; DB 10; Length 3013;
Best Local Similarity 59.3%; Pred. No. 0.0085;
Matches 127; Conservative 0; Mismatches 80; Indels 7; Gaps 1;
QY      1515  TTGAGCTTACGACCATATCATGTAATTGCTTCTCGATTAAGCTTTCATTA 1574
Db      1096  TTGAACCTTGAACCTTCCACATCATGTAATTGCTTGAAGATGAATAAGATTA 1155
QY      1575  TTCTCTTGCAAGCTGTGCTACTTACGAGAAGTC-----TGCTACAGACAGAA 1627
Db      1156  TTCTCTTAGTGAAGACGACGAATCTGCTTCTTAGAGGACACACAGCTGACATCA 1215
QY      1628  TGGCATTTTAACGGTACTTGGAAAAATGAGAAATGAGAAATGAGAAATTCATG 1687
Db      1216  TGGCATTTGATGGCACTTGGAAAGTAGACCGAAATGAGAAATGAGAAATTCAT 1275
QY      1688  CAATGGTAAAGCTTACTTTTGAATGCTTCT 1721
Db      1276  AATGGTAAAGGCTGCTTCTGCTATTTGCT 1309

RESULT 7
LOCUS   AC130496
DEFINITION
Rattus norvegicus clone CH230-4G2. *** SEQUENCING IN PROGRESS ***
ACCESSION
AC130496
VERSION
AC130496.4 GI:24818159
KEYWORDS
HTGS, HTGS_PHASE1, HTGS_DRAFT, HTGS_ENRICHED.
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 230897)
Wuzny,D.,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebech,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benham,F.,
Biswal,O., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carrroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinb,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabiela,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,M.,
Gunnarane,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

```

Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kwie, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, D., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, D., Lorenz, L., Louie, L., Louie, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapus, P., Martin, K., Martin, R., Martinez, E., McWhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Mlosoajevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunolu, G., Olarunpasegun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfennig, C., Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quirio, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shamsan, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Swatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, E., Uman, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 230897)
Worley, K.C.
Direct Submission
Submitted (11-VUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 230897)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23101330.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GBMN
Center clone name: CH230-4G2
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 210222 bases at least Q40
Consensus quality: 213326 bases at least Q30
Consensus quality: 215515 bases at least Q20
Estimated insert size: 218065; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 157719: contig of 157719 bp in length
157720 157819: gap of unknown length
157820 230897: contig of 73078 bp in length.
Location/Qualifiers
1. 230897
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-4G2"
1. 2206
/note="wgs end-extension
clone_end:Sp6
complement(72249..73061)
/note="clone boundary
clone_end:Sp6
site:BCORI
end sequence:BH306439"
157820..159219
/note="wgs contig"
182442..183960
/note="wgs_contig"

ORIGIN
Query Match 2.6%; Score 62; DB 2; Length 230897;
Best Local Similarity 59.3%; Pred. No. 0.0045;
Matches 127; Conservative 0; Mismatches 80; Indels 7; Gaps 1;
QY 1515 TTGAGCTTAAAGCCAGCATCATCATGTAATTCCTTCCTGTAAGCCTGTCATTA 1574
156173 TTGAACCTTGAATTCACATTCATGTAATGTTGTAAGATTAAGAAATGAATTA 156232
QY 1575 TTCTCTTTCGAAAGCTCTCTACTTACCAAGATC-----TGCTACAGACGAAGA 1627
156233 TTCTCTGATGAGACAGACGACCAATCTCTGCTTCTTAAGAGACACAGCTGACATCA 156292
QY 1628 TGGCATTTAAAGCTACTTGGAAATAGAGAAATATGAGAACTATGAAATTCATGAG 1687
156293 TGGCATTTATGACACTTGGAAATGAGACCGGAATGAGAACTATGAAATTCATGAG 156352
QY 1688 CAATGGTAAGCCTTACTTTTGAATGCTTCT 1721
156353 AAATGGTAAGGCTGCTTCTGCTATTTGCT 156386
Db

RESULT 8
AX083744 1141 bp DNA linear PAT 28-FEB-2001
LOCUS AX083744
DEFINITION Sequence 22 from Patent WO0111061.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1 Kunst, L. and Clemens, S.
TITLE Regulation of embryonic transcription in plants
JOURNAL Patent: WO 0111061-A 22 15-FEB-2001;
UNIVERSITY OF BRITISH COLUMBIA (CA)
FEATURES
source
1..1141
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1..1141
promoter

ORIGIN	/note="consensus sequence of A.t., L.a., and B.n. FABI promoters"				
Query Match	2.44	Score 57.8	DB 6	Length 1141	
Best Local Similarity	10.9%	Pred. No. 0.077			
Matches	82	Conservative 263	Mismatches 407	Indels 0	Gaps 0
QY	1565	TGTTCTAAATTCCTTTGAAAGCTCTGCTACTTACCAAGAACTGCGCTACGAGAGAA	1624		
DB	201	DRMSBRKMYGMBWNNNSYDYTYWVDDCKRVRWVTRGRMYVYVAMEJAH	260		
QY	1625	AGATGCACTTAAACGCTACTTGAAGAAATAGAGAAATGAACTATGAAATTCATG	1684		
DB	261	RRVNNGTBAMAYRRVNTNNNNNNNAAMCKRATYWGNNRABVSTCTYWSKTYRTS	320		
QY	1685	AAGCAATGGTAAAGCTCTTTTGAATGCTCTTAAAGAGATACCATACGCGG	1744		
DB	321	WANNCRAGDANKHKWKMGAAGVYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	380		
QY	1745	GAATPACAACTTAAGCTGTCATGAACTACATCTGCTAAGCTCTGCTCTGCTG	1804		
DB	381	AAHYSRKMVBKRYKTVNNNNNGTTMKRWMAWYKMDMBGTNNNNNGRTYTGWT	440		
QY	1805	TATTTGCCCTTCACATTCCTGCACTTATTTGAAAGCTCTATAGAGGGAATAC	1864		
DB	441	NKMTWYTKKANKCKRABDHKTCHNNNTTWKMTYNNNNNNNNNNNNNNNNNNNNNN	500		
QY	1865	AAGAAAGAAAACATCTGATTTTATTTGATTCGCAATATTTATGCTAATTT	1924		
DB	501	TWYMMWRRYAHANNNNNDYMWKACITWKYBVCSSKNNNNYAAWYTKSSMYTSRYR	560		
QY	1925	CCAGTAGAGGCATTCACGAGAAATTTAATAGATATATGTAAGAAATATTTTGA	1984		
DB	561	NNSWRMSDTRSMGRANNYRABHYGKMTFRWBHSHTBHRAGAAYHMMEMBAK	620		
QY	1985	TAAAGCTGTTGAAAAATTACACAAGAGGAAATGCTGCTCCAGTTTTCAGCAAC	2044		
DB	621	HCKAWYKAKKYAGAGSSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	680		
QY	2045	ACATGATTTGAGCTATTTAATCATGCTAGTCTTACTTACTTACTTACTTACTT	2104		
DB	681	BAANNAYYTHANNWGCMMNATDRTYMKNNNNNNNNNNNNNNNNNNNNNNNNNNNN	740		
QY	2105	AATATGATGTAACATACTACTAGTCTAGTATGATGATGATGATGATGATGATG	2164		
DB	741	AVRAAKHMRANKMARGWADDAABTDRKNNGATTKTNTNNNNNTYRGVYNTTARD	800		
QY	2165	TGCTCAGTGAAGAGTGAGAGAGTGAAGTCTGAGTCAAGATTCGGCTAAGCTCC	2224		
DB	801	GMANN	860		
QY	2225	TCAACTACAGAAAAAGTCAATATAAAATGCAATGATGTTCTATTTGTTTCTCT	2284		
DB	861	WRBAYTNN	920		
QY	2285	GCTTGAATGTAATGATATATATATTTT	2316		
DB	921	NNNNNTKTYBHAAMNN	952		
RESULT 9					
MUSFABP1					
LOCUS					
DEFINITION	Mouse Fabp1 gene, exons 1-4.				
ACCESSION	M65033				
VERSION	M65033.1 GI:193218				
KEYWORDS					
SOURCE	Mus musculus (house mouse)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases)				

AUTHORS	Green,R.P., Cohn,S.M., Sacchettini,J.C., Jackson,K.E. and Gordon,J.I.				
TITLE	The mouse intestinal fatty acid binding protein gene: nucleotide sequence, pattern of developmental and regional expression, and proposed structure of its protein product				
JOURNAL	DNA Cell Biol. 11 (1), 31-41 (1992)				
MEDLINE	92153306				
PUBMED	1739433				
REFERENCE	2 (bases 1 to 5039)				
AUTHORS	Green,R.P., Sacchettini,J.C., Jackson,K.E., Cohn,S.M. and Gordon,J.I.				
TITLE	The mouse fatty acid binding protein gene (Fabp1): nucleotide sequence analysis and predictions of the structure of its protein product				
JOURNAL	Unpublished (1991)				
COMMENT	Original source text: Mus musculus (strain DBA/2J) Adult Liver DNA.				
FEATURES					
source	Location/Qualifiers				
	1..5039				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/strain="DBA/2J"				
	/db_xref="taxon:10090"				
	/issue_type="Liver"				
	/dev_stage="Adult"				
	467..547				
misc_feature	/note="Domain I: homology with human I-FABP promoter region"; putative"				
	524..537				
repeat_region	/note="one of two repeats homologous to consensus repeat found in number of other lipid binding proteins which are expressed in small intestine"; putative"				
	727..817				
misc_feature	/note="Domain II: homology to human I-FABP promoter region"; putative"				
	870..966				
misc_feature	/note="Domain III: homology with human I-FABP promoter region"; putative"				
	884..897				
repeat_region	/note="Second of two repeats homologous to consensus repeat"; putative"				
	937..942				
TATA_signal	/note="putative"				
	967..1069				
exon	/note="putative"				
	1/label=Exon1				
exon	2326..2497				
	/note="putative"				
	1/label=Exon2				
exon	4291..4398				
	/note="putative"				
	1/label=Exon3				
exon	4818..5039				
	/note="putative"				
	1/label=Exon4				
ORIGIN					
Query Match	2.44	Score 56.6	DB 10	Length 5039	
Best Local Similarity	58.5%	Pred. No. 0.11			
Matches	121	Conservative 0	Mismatches 79	Indels 7	Gaps 1
QY	1516	TGAGCTTTAGCCAGACATCATGTAATGCTTCTGATAGAGCTGTTCAAT	1575		
DB	884	TGAACCTTGAACCTTCCACATCATGATATGTTGTTCAAGATGAATGAAT	943		
QY	1576	TCTCTTGGAAAGCTCTGCTACTTACAGAGTC-----TGCCTACAGACAGAAAGAT	1628		
DB	944	TCTCTAGTGAACAGACTGACCTGCTTCTCTAGAGACACACAGCTGAGTCAAT	1003		
QY	1629	GGCATTTACGCTACTTGAAGAAATAGAGAAATGGAATGAAATTCATGGAAC	1688		
DB	1004	GGCATTCAGCGCACTGGAAGATGACCGAAGCAACTATGAAAGTTATGAGGAA	1063		
QY	1689	AATGGTAAGCTTACTTTTGAATG	1715		

[illegible]

CDS
complement(join(10868..10972,11088..11177,11367..11617,
11727..11807,11991..12324))
/gene="PPD0335c"
/note="Similar to Plasmodium falciparum hypothetical 33.9
kDa protein mal4p2.04 SMALL:090U08 (EMBL:AL035475) (286
aa) fasta scores: E(): 4.9e-117, 98.25% id in 286 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD49160.1"
/db_xref="GI:23498189"
/db_xref="UniProt/TREMBL:O811W3"
/translation="MFNKLVSRTSYFLYDTRVNLNKLKLYKQMDYLKRMINY
FNNRKYNTDVIDEAKCWEKFKKELIQANGYIPALIMKXIEERCIQHKLIDEAF
VEEDGHSILFLFRVHIGKVCNVSVHSDPIEGHIFELKARAVENQITIN
PESIGLIGCPAYINGHYOLAMNSIKGVGSANI RGPFOIDVSKLTYPENYSIKL
KOLLVLPDQGNVIFSEBYNLDDETLVWTEPERKIMETPLPDDYDNPVNVKGRKIQ
LYTKDYWPKO"
CDS
complement(join(12996..13021,13108..13198,13312..13425,
13569..18095,18336..18767,18885..18910,19003..19037,
19326..19336,19432..19509))
/gene="PPD0335c"
/note="Alternative start site possible; gene model
slightly modified from submitted version (SMALL:090U07
EMBL:AL035475) Plasmodium falciparum hypothetical 205.8
kDa protein sequence
Similar to Plasmodium falciparum hypothetical 205.8 kDa
protein mal4p2.05 SMALL:090U07 (EMBL:AL035475) (1734 aa)
fasta scores: E(): 0, 99.25% id in 1742 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD49161.1"
/db_xref="GI:23498190"
/db_xref="UniProt/TREMBL:O811W2"
/translation="NNPFLITLLFPFIMFLKIAMGVHFGMCDSPLEFENRIS
EDRTYCKSKSEKYNLDLIDLYNDICTETKLEEDDESCKGFKCQILGTTTHAKN
KVALERKFEEDYLSQFVYKYNNDKNYINNDSBCGVNNISFDNENPYFNKYD
KNKFINPLEYKYLKGEVYITPGNVLHITNGQCNLIEKDILKLVSRNKKKY
KYCMNDVSYNS INKDRNNLNNVNLKRLKRTKYDKNKCYKFRSSVNSGKREK
ALSKAKKREKLMKNNKKRKEKESLNVEMENLNCYTKSKNINVIENNLK
NEKLEKIQKVNHDYSLOKLRKKKNDSTIKIEIRKKNRITNDKDKDKQIQ
NVSLSHLEKSLKNDKSEFLCSNMBIPVNNREKCEKIKERQTLFGSNTNTDLYK
DOISNKEDSLKNNIEINNESGVNIKKITRSYSNDIKSDYKKNKNTQVNDY
WYVNDYSOSCKEKREKRSYNNNDYSOSCKEKRSYNNNDYNNNDYNNIVE
YNNVSHDYSYNNKEGNIIIPYNDITEENSIIENKNDLDTOMEOLOSKEIKES
NEDDKDNTFTNLSHSIKERIKODINDNNEILNPLENNKNTSHVVEKK
NPSKKNKVDNRINKKDNQIININDIPENIPLMLILOHEKTKNNKKEKREKY
NIVBELGLPANCVNKLKALKEKRLPSRIEKLIDVCMYKKNONICFRKNKMTK
NNIVSDDEKRYGPIVSSRSYIPFPNNDKNNIHIKKYNNRYLLNFDGKNDIYN
NNGEBEKKNVPLPIKPLANTINVAKEVNKKNNRIKCYKKKCNKDNKINSIGIYD
KALNMGIKSYLKKNNIMVSEKMSIKDOLLKYNNMYSHLKNKCKITMSLSC
NMOKSDNVCKNNIVKYVNNHNAQNNKEASAIKKYNIHGHKNGIINRSTHS
DYKVDISONGLSIDIKDNKEVILISVLFPMNTHKEGMYTKRSIMRNNGLEKIKY
KKRBGNSFTCEIKNNDCQVATNENKGAJDSLNOEILIEKFPLOJENRBERIE
EEKQICLVDRKNNINNVDEKRSIDHPSYRVLKMGSSNNEGYSNTOKILNENKE
KNVNEKKGENDENKKEKENDENKKNENKENDENKENDENNNNSYFVNS
DTEFLCNLSNLIPINNKNSILIPSENEKGIISQKEBQNI SPKINNKKDOLCKIN
ESDYSDKOYSLVNLSEIKIKYKSSNSKIRGIEKKKINDDYVDLKNINRSTLFF
LTKKYLKSSRLIINBEDCONNVYKKKKBAKKKLNKINVINPINSIIEENSS
EYNPVKKKNNCWNKFTETRSKSIILSEIFAVKKNKRNATNLRSEBQFSSIGLVEK
GENKRIIEEKDEBIYEKIKNNKBNFPNNLTQLLPFKSAENINTSGSFTEKIRHVX

Query Match 2.2% Score 51.8; DB 3; Length 234081;
Best Local Similarity 45.6%; Pred. No. 0.65;
Matches 222; Conservative 0; Mismatches 262; Indels 3; Gaps 1;

DB 1834 TATTGAAAAAGCTATAGAGGGAATACAAGAAAGAAAACATTCTGATTATTG 1893
156675 TAT 156734
QY 1894 CATTGCATATCTTATGCACTTATGCTAATTCAGTAGAGGCACTTCAGACAGAAATTAA 1953
DB 156735 TATACAT 156794

QY 1954 ATAGATATATATGTAAGGAATATATATTTTGTATAGACTGTTGAAAAATTACACAAGAG 2013
DB 156795 ATACGATTTAT 156854
QY 2014 GAAATTCGTGCTCCGACGTTTTCGAAACACACATGATTTGATCATTTTAAACATGCTNG 2073
DB 156855 ACACATTTTTTTTTT---TTTTTAAAGTATATATATATATATATATATATATATAT 156911
QY 2074 TCGTACTTAAAGCTTGACAAACGCGCTGATATATGATGTAACATATACCTTACT 2133
DB 156912 TATATTAATTAATGAATGATATATATATATATATATATATATATATATATATATAT 156971
QY 2134 TCGATATGATTTGATTTATTAACAGCTGAACACTGCTCAGTGAAGTGAGAGAGTAA 2193
DB 156972 TGTGTTATGTTTATATGAGAAATATATATATATATATATATATATATATATATAT 157031
QY 2194 ACTGAGTCAGAAATTCGCGCTTAAGCTCCCTCAACTACAGAAAAAGTCACATATAAAT 2253
DB 157032 GATATTTATTTACATGATGTAATATATATATATATATATATATATATATATATAT 157091
QY 2254 GGAACATGATGTTGATTTGTTTCTGCTGCTGATGTTAATGATTTATATATTTT 2313
DB 157092 TTAAT 157151
QY 2314 TTTTCTT 2320
DB 157152 TTTTTTT 157158

RESULT 11
RATFABPX 564 bp mRNA linear ROD 27-APR-1993
LOCUS Rat intestinal fatty acid binding protein (FABP) mRNA, complete
DEFINITION cds
ACCESSION M35992
VERSION M35992.1 GI:204087
KEYWORDS fatty acid binding protein.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 564)

REFERENCE 1 (bases 1 to 564)
AUTHORS Gordon, J.I. and Lowe, J.B.
TITLE Analyzing the structures, functions and evolution of two abundant
gastrointestinal fatty acid binding proteins with recombinant DNA
and computational techniques
JOURNAL Chem. Phys. Lipids 38 (1-2), 137-158 (1985)
MEDLINE 86053743
PUBMED 3840724
COMMENT Original source text: Rat adult small intestinal epithelium, cDNA
to mRNA.

FEATURES
source location/Qualifiers
1..564
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
13..411
/note="fatty acid binding protein (FABP)"
CDS
13..411
/codon_start=1
/protein_id="AAA41141.1"
/db_xref="GI:204088"
/translation="MAFDGTVVDNENYKEMKNGINVVRKLGADNLTLLTIOE
GNKFTYKESNFRNIDIVVELGVDRAYSADGTSLGTWTMGNTLVGKFRVDNKE
LIAVRISGNEILIQTYTEGVAKRIFKE"

Query Match 2.1% Score 50; DB 10; Length 564;
Best Local Similarity 75.6%; Pred. No. 3.9;
Matches 62; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1615 ACAGACAGAAAGATGCACTTAAACGTACTTGAAAATGAGAAAAATGAGAACTATGAA 1674

Db 1 ACAGTCATCATCATGATTTGATGCACTTGAGAAAGTAGACCGGAATGAGAACTATGAA 60
QY 1675 AAATTCATGAGCAATGGCTA 1696
Db 61 AAGTTCATGAGAAATGGGCA 82

RESULT 12
XELIFABP 619 bp mRNA linear VRT 16-MAR-1994
LOCUS Xenopus laevis intestinal fatty acid binding protein mRNA, complete cds.
DEFINITION

ACCESSION 119946 GI:311053
VERSION 119946.1
KEYWORDS fatty acid; fatty acid binding protein; hormone-dependent regulation; metamorphosis; thyroid hormone; thyroid hormone dependent.

SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; 1 (bases 1 to 619)
REFERENCE Shi, Y.B. and Hayes, W.P. Thyroid hormone-dependent regulation of the intestinal fatty acid-binding protein gene during amphibian metamorphosis Dev. Biol. 161 (1), 48-58 (1994)

JOURNAL MEDLINE 94123868
PubMed 8293885
REFERENCE 2 (bases 1 to 619)
AUTHORS Shi, Y.B.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1993) Yun-Bo Shi, Laboratory of Molecular Embryology, Building 6, Rm. B1A02, NICHD/NIH, Bethesda, MD 20892, USA

COMMENT Original source text: Xenopus laevis cDNA to mRNA.
FEATURES
source location/Qualifiers
1..619
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/cell_type="epithelial cell"
/tissue_type="intestine"
/dev_stage="tadpole"
/tissue_1lb="Lambda Zap"
41..439
/codon_start=1
/product="fatty acid binding protein"
/protein_id="AAC38012.1"
/db_xref="GI:311054"
/translation="MADGTWKVDRESEYKEMVWGCVNIIVKRIKAHNLKVTIOQ GNNFTVKSSTFNIEIKFTLAPFPYSIADGELGAWFLQDNOLLGFTFRDNGKV LQTRROIIGDELQVTEYEETESKRIFKRG"

ORIGIN

Query Match 2.1%; Score 49.8; DB 5; Length 619;
Best Local Similarity 68.3%; Pred. No. 4.2;
Matches 69; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1595 TACTTACGAGAGTCTGCTACAGACAGAAAGATGATTACCGGACTTGGAAAATG 1654
Db 9 TGGCTAGCCAAAGCTCTCTCCACCAATCATCTAGGCTTGTATGGAAGGTTG 68

QY 1655 AGAAAAATGAGAACTATGAAAAATTCATGGAAGCAATGGGT 1695
Db 69 ACAGAACTGAAACTATGAGAAATTCATGGAAGTATGGGT 109

RESULT 13
LOCUS AX771590 5204 bp DNA linear PAT 02-JUL-2003
DEFINITION Sequence 303 from Patent WO03004646.
ACCESSION AX771590

VERSION AX771590.1 GI:32438388
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
O'Mahony, D.J., Byrne, D., Brayden, D., Lambkin, I. and Higgins, L. Genetic analysis of Peyer's patches and M cells and methods and compositions targeting Peyer's patches and M cell receptors Patent: WO 03004646-A 303 16-JAN-2003;
ELAN CORPORATION, PLC (IE)

JOURNAL LOCATION/Qualifiers
FEATURES
source 1..5204
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 2.1%; Score 49.8; DB 6; Length 5204;
Best Local Similarity 59.2%; Pred. No. 3.1;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCAATCATCATGTAATGCTTCCGTATAGCCGTTCATPAAT 1575
Db 971 TGAATTTTAACTTCCACATCACAGTATGAGTTGGTTCAAGTAAAGAAATTAATPAAT 1030

QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGAGTGCCTACAGAGAAGA 1627
Db 1031 TCTCGCCCAAGAGACACACCGTAATCTTGCTGCTCAGAGGCTGACT-CACTGAATCA 1089

QY 1628 TGGCATTTAAAGGTAATTTGGAAATAGAGAAATAGAACTATGAAAAATTCATGAG 1687
Db 1090 TGGCGTTTACACGACCTTGGAAAGTAGACCGGAGTAAAAAATATGACAGTTCATGAAA 1149

QY 1688 CAATGGTAAGCCTTCTTTTGAATGCT 1718
Db 1150 AATGGTAAGACCTTATTTCTTTGTGGCT 1180

RESULT 14
HUMFABP 5204 bp DNA linear PRI 08-NOV-1994
LOCUS Human, intestinal fatty acid binding protein gene, complete cds, and an Alu repetitive element.
DEFINITION M18079 J03465
ACCESSION M18079.1 GI:182351
VERSION Alu repeat; fatty acid binding protein.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 5204)
Sweetser, D.A., Birkenmeier, E.H., Kisek, I.J., Zollman, S., Sparkes, R.S., Mohandas, T., Lusis, A.J. and Gordon, J.I. The human and rodent intestinal fatty acid binding protein genes. A comparative analysis of their structure, expression, and linkage relationships J. Biol. Chem. 262 (33), 16060-16071 (1987)

JOURNAL Original source text: Human DNA (library of T.Maniasis), clone lambda-HIFABP.
MEDLINE 88058967
PubMed 2824476
COMMENT Draft entry and computer-readable sequence for [1] kindly provided by D.Sweetser, 19-JAN-1988.

FEATURES
source 1..5204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="4q28-q31"
prim_transcript 1028..54393
/note="FABP1 mRNA (alc.) and introns"

	prim_transcript	1053..>4393	/note="FABP1 mRNA (alt.) and introns"
CDS		join(1089..1155,2350..2522,3546..3653,4098..4148)	
	exon	/note="fatty acid binding protein"	
		/codon_start=1	
		/protein_id="AA52417.1"	
		/db_xref="GI:182352"	
gene		/translation="MAPDSTMKYDRSENYDKFMKGVNIYKRLLAHDLKLTIOE GNKTIVKSSAFENIEVFEGLGTFENLMDIELRGTWLBEGNKLIGFKRTDNGNE LNTVREIIGDELVTQTVYEVEAKRIFFKO"	
	exon	<1089..1155	
		/gene="FABP2"	
		<1089..1155	
		/gene="FABP2"	
		/note="fatty acid binding protein; G00-119-127"	
	intron	/number=1	
		1156..2349	
		/note="FABP1 intron A"	
	exon	2350..2522	
		/number=2	
		2523..3545	
		/note="FABP1 intron B"	
	exon	3546..3653	
		/number=3	
		3654..4097	
	intron	/note="FABP1 intron C"	
		4098..>4148	
	exon	/note="fatty acid binding protein"	
		/number=4	
		4466..4742	
	repeat_region	/note="Alu repeat"	
		4466..4472	
	repeat_region	/note="5' direct repeat"	
		4736..4742	
	repeat_region	/note="3' direct repeat"	
ORIGIN	156 bp upstream of HindIII site; chromosome 4q28-q31.		
Query Match	2.1%; Score 49.8; DB 9; Length 5204;		
Best Local Similarity	59.2%; Pred. No. 3.1;		
Matches 125; Conservative	0; Mismatches 77; Indels 9; Gaps 2;		
Oy	1516 TGAGCTTACGACGCATCATCATGAATTGGTTCCTCGATAAGCGTTCATTAAT	1575	
Db	971 TGAAGCTTAACTTCACATCAAGTAGAAGTGTTGCAAGATAAGAAATATATAAT	1030	
Oy	1576 TCCTC-----TTGCAAAGCTCGACTTACACAGAAGTCGCTACAGACGAAGA	1627	
Db	1031 TCTGCCCAAGACAGACTGAATCTCTAGCTGCTAGAGGCTACT-CAACTGAAATCA	1089	
Oy	1628 TGGCATTACGCTACTTGGAAAAATAGAAAAATGACAACCTATGAAAATTCATGGAAG	1687	
Db	1090 TGGCGTTTGAACAGCACTTGGAAGTAGACCAGAGTAACTATGACAAAGTTCAATGGA	1149	
Oy	1688 CAATGGGTAAAGCTTACTTTTGAATGCTT	1718	
Db	1150 AAATGGTAAAGACTTATTTCTTGTGCT	1180	
RESULT 15			
LOCUS	AC008077c	200000 bp DNA linear HTG 20-JUL-1999	
DEFINITION	Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***, 20		
ACCESSION	AC008077		
VERSION	AC008077.1 GI:5540103		
KEYWORDS	HTG; HTGS_PHASE1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 200000) Stone,N.E., Schmutz,J.V., Cox,D.R. and Myers,R.M. Direct Submission		

REFERENCE	JOURNAL	Unpublished
2 (bases 1 to 200000)		
Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.		
AUTHORS		
TITLE	Direct Submmission	
JOURNAL	Submitted (20-JUL-1999) Department of Genetics, Stanford Human Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA	
COMMENT	NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preereved.	
	1 1146: contig of 1146 bp in length	
	1147 10118: gap of unknown length	
	10119 11308: contig of 1190 bp in length	
	11309 20280: gap of unknown length	
	20281 21488: contig of 1208 bp in length	
	21489 30460: gap of unknown length	
	30461 31634: contig of 1174 bp in length	
	31635 40606: gap of unknown length	
	40607 41979: contig of 1373 bp in length	
	41980 50952: gap of unknown length	
	50952 52049: contig of 1098 bp in length	
	52050 61021: gap of unknown length	
	61022 62105: contig of 1084 bp in length	
	62106 71077: gap of unknown length	
	71078 72391: contig of 1314 bp in length	
	72392 81363: gap of unknown length	
	81364 82567: contig of 1203 bp in length	
	82567 91538: gap of unknown length	
	91539 92900: contig of 1362 bp in length	
	92901 101871: gap of unknown length	
	101872 103074: contig of 1203 bp in length	
	103075 112045: gap of unknown length	
	112046 113302: contig of 1257 bp in length	
	113303 122273: gap of unknown length	
	122274 123778: contig of 1505 bp in length	
	123779 132749: gap of unknown length	
	132750 134818: contig of 2069 bp in length	
	134819 143789: gap of unknown length	
	143790 145506: contig of 1717 bp in length	
	145507 154477: gap of unknown length	
	154478 156731: contig of 2254 bp in length	
	156732 165702: gap of unknown length	
	165703 167355: contig of 1653 bp in length	
	167356 176326: gap of unknown length	
	176327 178351: contig of 2025 bp in length	
	178352 187322: gap of unknown length	
	187323 188990: contig of 1668 bp in length	
	188991 197961: gap of unknown length	
	197962 200000: contig of 2039 bp in length.	
FEATURES	Location/Qualifiers	
source	1. 200000	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="4"	
ORIGIN		
Query Match	2.14; Score 49.8; DB 2; Length 200000;	
Best Local Similarity	59.24; Pred. No. 1.8;	
Matches 125; Conservative	0; Mismatches 77; Indels 9; Gaps 2;	
QY	1516 TGACCTTACCGACGACATCATCATGTAATTCCTTCTGTAAGCCTGTTCAATAAT	1575
DB	10910 TGAACCTTAGGCTTCCACATCATCAGTATGAAAGTTGGTTCAGATGAAGAATTAATAAT	10851
QY	1576 TCTC-----TTTGAAGCTCTGCTACTTACCGAAGCTTCGCTTAAGACAGAAAGA	1627
DB	10850 TCTGCGCCAGACGACGACCTGAAATCTCTAGCTGCGCTTAAGAGGCTGACT-CAACTGAAATCA	107922

QY 1628 TGGCATTTAACGGTACTTGGAAAAATAGAAAAATGAACTATGAAAAATTCATGGAG 1687
 Db 10791 TGGCGTTTGACAGCACTTGGAGGTAGACCGAGGTGAAAACTATGACAGTTCTATGSA 10732
 QY 1688 CAATGGGTAAAGCCTTACTTTTGTGATGCT 1718
 Db 10731 AAATGGGTAAAGACTTATTCTTGTGGCT 10701

Search completed: November 21, 2005, 00:03:22
 Job time : 9980 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2005, 15:37:53 ; Search time 1191 Seconds
(without alignments)
11834.512 Million cell updates/sec

Title: US-10-099-663-1

Sequence: 2381
1 agcttcctcgccgagaaagg.....atctgaagctcactatctcag 2381

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2381	100.0	2381	11	ADL90127	Adl90127 Chicken i
2	336	14.1	236	11	ADL90128	Adl90128 Chicken i
3	50.6	2.1	4590	5	Aah24065	Aah24065 Yeast AD
4	49.8	2.1	5204	6	ABK83946	Abk83946 Human CDN
5	49.8	2.1	5204	12	ADJ74588	Adj74588 Human int
6	49.8	2.1	5204	12	ADO78215	Ado78215 Human fat
7	49.2	2.0	2000	8	ADA71938	Ada71938 Rice gene
8	47.2	2.0	337	8	ABX40871	Abx40871 Bovine ES
9	47	2.0	2636	4	ABL15880	Ab115880 Drosophila
10	46.8	2.0	662	10	ADD29587	Add29587 Mouse bre
11	46.6	2.0	610	4	AAL19953	Aal19953 Human bre
12	46.6	2.0	9742	6	ABL70479	Ab170479 Chemicall
13	46.2	1.9	6222	6	ABL32692	Ab132692 Human imm
14	45	1.9	10652	6	ABN80056	Abn80056 Human che
15	44.4	1.9	2000	8	ADA71938	Ada71938 Rice gene
16	44.2	1.9	515	3	AAAI6469	Aaa16469 Human col
17	44.2	1.9	2252	13	ADQ39082	Adq39082 Human SNP
18	43.6	1.8	14023	6	ABL34104	Ab134104 Human imm
19	43.2	1.8	7143	6	ABL32983	Ab132983 Human imm
20	42.8	1.8	51289	13	ABD33486	Abd33486 Human can

C	21	42.6	1.8	513	12	ACH79811	Ach79811 Human gen
C	22	42.6	1.8	611	3	AAA16371	Aaa16371 Human col
C	23	42.6	1.8	2844	10	ABT31903	Abt31903 Human bre
C	24	42.6	1.8	5501	6	ABT32055	Abt32055 Human imm
C	25	42.4	1.8	230	2	AAAX1681	Aax11681 Human bia
C	26	42.4	1.8	230	2	AAAX1293	Aax1293 Human bia
C	27	42.4	1.8	452	3	AAAT7858	Aaa7858 cDNA enco
C	28	42.4	1.8	452	4	AAAT8596	Aaa8596 Colon tum
C	29	42.4	1.8	452	8	ABZ32782	Abz32782 Human col
C	30	42.4	1.8	461	4	AAL10372	Aal10372 Human bre
C	31	42.4	1.8	682	11	ACN80571	Acn80571 Breast ca
C	32	42.4	1.8	5908	4	AA545387	Aaa545387 Chemicall
C	33	42.4	1.8	5908	6	ABK28322	Abk28322 DNA trans
C	34	42.2	1.8	453	6	ABE7809	Abt67809 Oesophagu
C	35	42.2	1.8	453	6	ABE62188	Abt62188 Colon ade
C	36	42.2	1.8	6621	6	ABJ32912	Abj32912 Human imm
C	37	42	1.8	9095	6	ABO67061	Abog67061 Human ang
C	38	41.8	1.8	10048	6	ABL70313	Ab170313 Chemicall
C	39	41.8	1.8	10048	6	AA561251	Aaa61251 Human gen
C	40	41.8	1.8	137870	10	ADG89426	Adg89426 Human mat
C	41	41.6	1.7	900	4	AAH34536	Aah34536 Human col
C	42	41.6	1.7	3639	6	ABE62356	Abt62356 Colon ade
C	43	41.6	1.7	3639	12	ADJ74884	Adj74884 Marker ge
C	44	41.6	1.7	3719	9	ADA10896	Ada10896 Human CDN
C	45	41.4	1.7	447	8	ABX50540	Abx50540 Bovine ES

ALIGNMENTS

RESULT 1	ADL90127	ADL90127 standard; DNA; 2381 BP.
ID	ADL90127	
AC	ADL90127;	
XX		
DT	20-MAY-2004	(first entry)
XX		
DE	Chicken intestinal fatty acid binding protein, iFABP, gene, 5' region.	
XX		
KW	Chicken; ds; intestinal fatty acid binding protein; iFABP;	
XX	gut specific promoter; transgenic.	
OS	Gallus gallus.	
XX		
PN	US2003177516-A1.	
XX		
PD	18-SEP-2003.	
XX		
PF	14-MAR-2002; 2002US-00099663.	
XX		
PR	14-MAR-2002; 2002US-00099663.	
XX		
PA	(HORS/) HORSEMAN N D.	
XX		
PI	(PRAT/) PRATT S L.	
XX		
PI	Horseman ND, Pratt SL;	
XX		
DR	WPI; 2003-898653/82.	
XX		
PT	New nucleic acid molecule comprising an isolated avian gut-specific gene	
XX	expression control region, useful for regulating heterologous nucleic	
PT	acids in transgenic avians, and for generating transgenic birds.	
PS	Claim 1, SEQ ID NO 1; 28bp; English.	
XX		
CC	The invention relates to an isolated nucleic acid comprising an isolated	
CC	avian gut-specific gene expression control region appearing as	
CC	ADL90127(Chicken intestinal fatty acid binding protein, iFABP, gene, 5'	
CC	region or ADL90128 (Chicken iFABP promoter) or its degenerate variant.	
CC	Also included are a recombinant DNA molecule comprising an isolate avian	
CC	gut-specific gene expression control region operably linked to a nucleic	
CC	acid insert encoding a polypeptide, an expression vector that integrates	

into a host cell (and comprising the isolated avian gnt-specific gene expression control region), expressing a heterologous polypeptide in a host cell (by transfecting a eukaryotic cell with the recombinant DNA molecule, and culturing the transfected cell in a medium suitable for expression of a heterologous polypeptide under the control of an avian intestinal fatty acid binding protein (IFABP) or cp15 gene expression control region encoded by the recombinant DNA molecule), a eukaryotic cell transformed with the expression vector (or its progeny, which expresses a heterologous polypeptide) and a transgenic avian having a heterologous polynucleotide sequence comprising the nucleic acid insert. The nucleic acids are useful for regulating heterologous nucleic acids in transgenic avians, as probes in nucleic acid hybridization assays for detecting the IFABP gene expression control region, and for generating transgenic birds. The present sequence is the Chicken intestinal fatty acid binding protein, IFABP, gene, 5' region.

XX Sequence 2381 BP; 775 A; 428 C; 469 G; 709 T; 0 U; 0 Other;

Query Match 100.0%; Score 2381; DB 11; Length 2381;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AGCTTCCTGGCGAGAAAAGGCTGTGGGTTCTTGTCCCTCACACAGCTTAAAGCAAAATC 60
DB 1 AGCTTCCTGGCGAGAAAAGGCTGTGGGTTCTTGTCCCTCACACAGCTTAAAGCAAAATC 60
QY 61 CCCAAGTTCAAAAAGTCCGCTGTGAAGAGATGCGTCACTTCAATGAAGTAATTANG 120
DB 61 CCCAAGTTCAAAAAGTCCGCTGTGAAGAGATGCGTCACTTCAATGAAGTAATTANG 120
QY 121 AATATATCATTAACAGAGCTCTGTGGCAGATCAGAGATTAACCTGTGGGCAAAATTT 180
DB 121 AATATATCATTAACAGAGCTCTGTGGCAGATCAGAGATTAACCTGTGGGCAAAATTT 180
QY 181 CTTAAAGTGTATAGTGAACAGAGAGTGTTCGCACTAAATGAGTAAATGTATACACA 240
DB 181 CTTAAAGTGTATAGTGAACAGAGAGTGTTCGCACTAAATGAGTAAATGTATACACA 240
QY 241 TTTGATCTTCTAGGAGACAAAAGGCTGTGAAAAACAATTAATTTCTGGTGCACAGTCACT 300
DB 241 TTTGATCTTCTAGGAGACAAAAGGCTGTGAAAAACAATTAATTTCTGGTGCACAGTCACT 300
QY 301 AGCAGCTGTCTTGGGTGCACTACAGCAACTTTGTTTGCAACATTAACAATCTTAAGTTG 360
DB 301 AGCAGCTGTCTTGGGTGCACTACAGCAACTTTGTTTGCAACATTAACAATCTTAAGTTG 360
QY 361 TTTTCTTTTCTCTTTCTTAACTTCTGTACAGTCTTAAAGTGAAGTAGTATTTAGT 420
DB 361 TTTTCTTTTCTCTTTCTTAACTTCTGTACAGTCTTAAAGTGAAGTAGTATTTAGT 420
QY 421 TACTTCCCTCTGATCCTCTTACCCAGATTAAGCATTTGCAAAATGAACTGAGTGA 480
DB 421 TACTTCCCTCTGATCCTCTTACCCAGATTAAGCATTTGCAAAATGAACTGAGTGA 480
QY 481 ATGGAAGACACACTATTTTGTATCACAGCAAAAGTCTAAATTTATAGTTATCTTCA 540
DB 481 ATGGAAGACACACTATTTTGTATCACAGCAAAAGTCTAAATTTATAGTTATCTTCA 540
QY 541 GTAAAACTTTTCTGAGGTCTGAGAAAGAAAAGAAATTAATGATTAACACAGACTAGTA 600
DB 541 GTAAAACTTTTCTGAGGTCTGAGAAAGAAAAGAAATTAATGATTAACACAGACTAGTA 600
QY 601 AATTTCTTAGTTAGAGCCAACTCTGTATCTGTGTATGAACAATTTATTTCAAGAT 660
DB 601 AATTTCTTAGTTAGAGCCAACTCTGTATCTGTGTATGAACAATTTATTTCAAGAT 660
QY 661 TCAGGATTTACATTTTGAAGCTAATGACAGAGATTTGTGCGCTTCATAGAAACAGA 720
DB 661 TCAGGATTTACATTTTGAAGCTAATGACAGAGATTTGTGCGCTTCATAGAAACAGA 720
QY 721 CTAACATAATCTGAGATTAGTACAAGAGATTTAGCACAGCAAAATTTGCTCAGTTTC 780
DB 721 CTAACATAATCTGAGATTAGTACAAGAGATTTAGCACAGCAAAATTTGCTCAGTTTC 780
```

```
QY 781 AAGTAGCACTATCTTGTGGGGAAGAAAGAGCTGAGCCAGTGTGCTCATTTTCGACAT 840
DB 781 AAGTAGCACTATCTTGTGGGGAAGAAAGAGCTGAGCCAGTGTGCTCATTTTCGACAT 840
QY 841 ATCTTCAACATTTTAAACCTGGGATCTATGGAATCAACAAGTGTGGTAAATTTCACT 900
DB 841 ATCTTCAACATTTTAAACCTGGGATCTATGGAATCAACAAGTGTGGTAAATTTCACT 900
QY 901 TAGCGACATCACTACTGTAAGTATGGAATGAGACAGAAACAGAGATTCAGTAATGGGCTAT 960
DB 901 TAGCGACATCACTACTGTAAGTATGGAATGAGACAGAAACAGAGATTCAGTAATGGGCTAT 960
QY 961 AATATAGGAATAGCTAGGAAGGTGCTGAATTTAGCTACTACTATTTAAAGATGAGGACA 1020
DB 961 AATATAGGAATAGCTAGGAAGGTGCTGAATTTAGCTACTACTATTTAAAGATGAGGACA 1020
QY 1021 CGAATGAGAAATATCATGCAATTTCTGTAGCTCAGCACTAGACTGAAAGTTTCTGAAA 1080
DB 1021 CGAATGAGAAATATCATGCAATTTCTGTAGCTCAGCACTAGACTGAAAGTTTCTGAAA 1080
QY 1081 CTGAACCGAGTTTCCCAAACCTACCTGTGAGATGTCAGTGGATCTTCAATCTCATGCTTAT 1140
DB 1081 CTGAACCGAGTTTCCCAAACCTACCTGTGAGATGTCAGTGGATCTTCAATCTCATGCTTAT 1140
QY 1141 TATGTGAGTGAATATGATTTCTACCAAAATTAAGATGACAAAGAGAGATTTGTGTTTT 1200
DB 1141 TATGTGAGTGAATATGATTTCTACCAAAATTAAGATGACAAAGAGAGATTTGTGTTTT 1200
QY 1201 ATCTGTGGGTAAATAGCTTTTCTCCAGTTGTATTAAGACCTCCACAGATTAATAGTC 1260
DB 1201 ATCTGTGGGTAAATAGCTTTTCTCCAGTTGTATTAAGACCTCCACAGATTAATAGTC 1260
QY 1261 CTATGCAACAAAGAAATGTCAATCATTTCTTATGCTCATTTATTTTATTATAGATA 1320
DB 1261 CTATGCAACAAAGAAATGTCAATCATTTCTTATGCTCATTTATTTTATTATAGATA 1320
QY 1321 GCCGGTTTTTATCTCAACTCAATTAAGTGAACGAATGAGGTTAGTACGTTTAA 1380
DB 1321 GCCGGTTTTTATCTCAACTCAATTAAGTGAACGAATGAGGTTAGTACGTTTAA 1380
QY 1381 TAAAGAGATTAATTAAGATATCATCATTTTGAAGCAATTAAGGAGGAGATTCAG 1440
DB 1381 TAAAGAGATTAATTAAGATATCATCATTTTGAAGCAATTAAGGAGGAGATTCAG 1440
QY 1441 CAAACAGTGTGCTTACAGTGAAGAAACAAGTTAAACTAAAGTGACCCCTCTTGACAA 1500
DB 1441 CAAACAGTGTGCTTACAGTGAAGAAACAAGTTAAACTAAAGTGACCCCTCTTGACAA 1500
QY 1501 GATCAATGCAAGTTGAGCTTTAGCGAGCAATCATATGATTAATGCTTCCGATTA 1560
DB 1501 GATCAATGCAAGTTGAGCTTTAGCGAGCAATCATATGATTAATGCTTCCGATTA 1560
QY 1561 AGCTGTTCATTAATTTCTTTGCAAAAGCTGTCTACTTACAGAAAGTCTGCTACAGAC 1620
DB 1561 AGCTGTTCATTAATTTCTTTGCAAAAGCTGTCTACTTACAGAAAGTCTGCTACAGAC 1620
QY 1621 AGAAGATGCAATTTAAACGTTACTTGAAAAATGAGAAAAATGAGAACTATCAAAAATTC 1680
DB 1621 AGAAGATGCAATTTAAACGTTACTTGAAAAATGAGAAAAATGAGAACTATCAAAAATTC 1680
QY 1681 ATGGAAGCAATGGGTGAAGCTTAATTTTGAATGCTCTTAAGAGAGATTAACCACTAC 1740
DB 1681 ATGGAAGCAATGGGTGAAGCTTAATTTTGAATGCTCTTAAGAGAGATTAACCACTAC 1740
QY 1741 GCGGGAATCAAACTTAAGCTGTTCATGAACTACCATCTGGCTAACTCTGCTTGTGT 1800
DB 1741 GCGGGAATCAAACTTAAGCTGTTCATGAACTACCATCTGGCTAACTCTGCTTGTGT 1800
QY 1801 CTGCTATTTTGCCTTTCACATTTGCCCTGCACTTATTTTGAAGAAGCTTATAGAGGGGA 1860
DB 1801 CTGCTATTTTGCCTTTCACATTTGCCCTGCACTTATTTTGAAGAAGCTTATAGAGGGGA 1860
```


FT	/tag= a	"Represented as * in the specification"
FT	/note= 3617	"Represented as * in the specification"
PT	/tag= b	"Represented as * in the specification"
FT	/note= 3649	"Represented as * in the specification"
FT	/tag= c	"Represented as * in the specification"
FT	/note= 3679	"Represented as * in the specification"
FT	/tag= d	"Represented as * in the specification"
FT	/note= 3819	"Represented as * in the specification"
FT	/tag= e	"Represented as * in the specification"
FT	/note= 3862	"Represented as * in the specification"
FT	/tag= f	"Represented as * in the specification"
FT	/note= 3864	"Represented as * in the specification"
FT	/tag= g	"Represented as * in the specification"
FT	/note= 3888	"Represented as * in the specification"
PT	/tag= h	"Represented as * in the specification"
FT	/note= 3890	"Represented as * in the specification"
FT	/tag= i	"Represented as * in the specification"
FT	/note= 3912	"Represented as * in the specification"
FT	/tag= j	"Represented as * in the specification"
FT	/note= 3914	"Represented as * in the specification"
FT	/tag= k	"Represented as * in the specification"
FT	/note= 3938	"Represented as * in the specification"
PT	/tag= l	"Represented as * in the specification"
FT	/note= 3939	"Represented as * in the specification"
FT	/tag= m	"Represented as * in the specification"
FT	/note= 3941	"Represented as * in the specification"
FT	/tag= o	"Represented as * in the specification"
FT	/note= 3943	"Represented as * in the specification"
FT	/tag= p	"Represented as * in the specification"
FT	/note= 4361	"Represented as * in the specification"
FT	/tag= q	"Represented as * in the specification"
XX	/note=	"Represented as * in the specification"
PN	WO200133977-A1.	
XX		
PD	17-MAY-2001.	
XX		
PX	06-NOV-2000; 2000MO-AU001362.	
PF		
XX	05-NOV-1999; 99AU-00003875.	
PR		
XX	(META-) METABOLIC PHARM LTD.	
PA		
XX		
PI	Belyea CI, Ng FM, Vaughan P;	
XX		
DR	WFI, 2001-328876/34.	
XX		
CC	The invention relates to novel transgenic organisms useful in the	
CC	production of functional food and drink products for the treatment or	
CC	prevention of obesity via the regulation of lipid metabolism. The	
CC	organisms comprise a polynucleotide encoding a growth hormone fragment	
CC	capable of stimulating the activity of hormone-sensitive lipase (the key	
CC	obesity and in the meat production industry.	
XX	Disclosure; Page 48-50; 54pp; English.	

[illegible]

DB 4262 RACRRNHHBDSHBSSTWRDCCSTMATSSCDTMDYSHSHSNWYSDNDS 4321
 QY 932 AGAAACAGACATCTACTGTAATGGCTATATATAGAAATACGTAGACGTCTCTGAA 991
 DB 4322 WNYSDMCRSTWRDRSDGRTTHACKCHTTDGDANBDDRTNRYRSHARRSBTABRAC 4381
 QY 992 TTTAGACTACTCTTTAAAGAGTGAAGACAGCAATGAGGAATATCATCGCAATTTCTGTAG 1051
 DB 4382 ATNASADMTADMNSTDMWDSSTWRDDCMNTCMNTSRCCNTMCDNTMANASADGSTSMARY 4441
 QY 1052 CTCAGCACTAGACTCTGAGAGTTTCTGAAACTGAACCGAGTTTCCCAACTACCTG 1106
 DB 4442 NRMATNNSTRATTCMANAGRRYWDCMNTSMARYNMAINRBSCTRTCAHBS 4496

RESULT 4
 ID ABR83946 standard; cDNA; 5204 BP.
 AC ABR83946;
 XX
 DT 14-AUG-2002 (first entry)
 DE Human cDNA differentially expressed in granulocytic cells #517.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US030821.
 XX
 PR 03-OCT-2000; 2000US-0237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.
 XX
 PS Claim 1; SEQ ID NO 517; 1149p; English.
 XX
 OS The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) GA by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile

CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease, also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;
 XX

Query Match 2.1%; Score 49.8; DB 6; Length 5204;
 Best Local Similarity 59.2%; Pred. No. 0 26; Mismatches 9; Gaps 2;
 Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCGACCATCATCATGTAATTCCTTCTGATPAAAGCTTTCATTAAT 1575
 DB 971 TGAACCTTTTAACTTCCACATCAAGATATGAAAGTTGTTCAAGAAATATATTAAT 1030
 QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCGAAAGTCTGCTTACGACAGAAAGA 1627
 DB 1031 TCTCGGCCAAGACACACCTGTAATCTCTAGCTGCTGAGAGGCTGACT-CAACTGAATATCA 1089
 QY 1628 TGGCATTTTAAAGGTAAGTGAATTAATGAGAAATAGGAAGTAAGTAATTCATGAGAG 1687
 DB 1090 TGGCGTTTGAACACATCTGGAAGGTAGACCGAGTAAATCTATGACCACTTATGAGAA 1149
 QY 1688 CAATGGGTAAGCCTTACTTTTGAATGCT 1718
 DB 1150 AATGGGTAAAGACTTATTTCTTGTGGCT 1180

RESULT 5
 ID ADJ74588 standard; DNA; 5204 BP.
 AC ADJ74588;
 XX
 DT 06-MAY-2004 (first entry)
 DE Human intestinal fatty acid binding protein gene SEQ ID NO:12.
 XX
 KW de; gene; human; intestinal fatty acid binding protein; restenosis;
 KW coronary angioplasty; balloon coronary angioplasty;
 KW stent coronary angioplasty.
 XX
 OS Homo sapiens.
 XX
 PN WO2004015104-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 20-MAR-2003; 2003WO-JP003478.
 XX
 PR 09-AUG-2002; 2002JP-00233041.
 XX
 PA (NAGO-) NAGOYA IND SCI RES INST.
 XX
 PI (GIFU-) GIFU INT INST BIOTECHNOLOGY.
 XX
 DR Yamada Y, Yokota M;
 XX
 DR WPI; 2004-180672/17.
 XX

PT Analysis of specific gene polymorphisms in clinical nucleic acid sample
PT for prediction of risk of restenosis after balloon or stent coronary
PT angioplasty.
PS Disclosure; SEQ ID NO 12; 164pp; Japanese.
XX
XX
XX The invention relates to a novel method for predicting the risk of
CC restenosis after coronary angioplasty comprising analysing specific gene
CC polymorphisms in a clinical nucleic acid sample. The method is useful for
CC the diagnosis of the genetic risk of restenosis following balloon or
CC stent coronary angioplasty. The method has high accuracy and high
CC estimation ratio. The present sequence is used in the exemplification of
CC the invention.
XX
SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;
Query Match 2.1%; Score 49.8; DB 12; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.26;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;
QY 1516 TGAGCTTTAGCCAGCAGCATCATCATGTAATGCTTCTGTAAGCCTTCATTAAT 1575
DB 971 TGAACCTTTAGCTTCCACATCACTACATGATGAGTTGGTTCAAGATTAATTAATTAAT 1030
QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGCTGCTTACAGACAGAAAG 1627
DB 1031 TCTGCCCAAGAGACAGACCTGAAATCTTACTGCTCTTAGAGGCTGACT-CAACTGAATCA 1089
QY 1628 TGGCATTTAACGGTACTTGGAAAAATAGAAAAATGAAACTTGTAAAAATTCATGGAAG 1687
DB 1090 TGGCGTTTACACGACTTGAAGGTAGACCGAGGTGAAAACTATGCAAGTTTCATGGAAA 1149
QY 1688 CAATGGTAAAGCCTTATTTTGAATGCT 1718
DB 1150 AATGGGTAAAGACTTATTCTTTGTGCT 1180
RESULT 6
ADO78215
ID ADO78215 standard; DNA; 5204 BP.
XX
XX ADO78215;
AC
DT 29-JUL-2004 (first entry)
XX
DE Human fatty acid binding protein 2 Alu repeat region DNA with SNP Seq 9.
XX
XX human; db; single nucleotide polymorphism; SNP; hereditary risk;
KM coronary spasm; tumour necrosis factor alpha;
KM NADH/NADPH oxidase p22 phox gene; apolipoprotein E; apolipoprotein C-II;
KM streptolysin 1; interleukin 6;
KM endothelium type nitrogen monoxide synthetase;
KM fatty acid binding protein 2; cardiac sudden death;
KM variant angina pectoris; myocardial infarction.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT variation /tag= a replace(2445,A)
FT /standard_name= "Single nucleotide polymorphism"
XX
XX JF2004113093-A.
XX
XX 15-APR-2004.
XX
XX 25-SEP-2002; 2002JP-00280031.
XX
XX 25-SEP-2002; 2002JP-00280031.
XX
XX (NAGO-) ZH NAGoya SANGYO KAGAKU KENKYUSHO.
PA (GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUO.
XX

DR WPI; 2004-360127/34.
XX
XX
PT Detecting genotype in nucleic acid sample useful for diagnosing risk of
PT coronary spasm, by analyzing polymorphisms in tumor necrosis factor alpha
PT gene, NADH/NADPH oxidase p22 phox gene and apolipoprotein E gene, in
PT nucleic acid sample.
PS Disclosure; SEQ ID NO 9; 59pp; Japanese.
XX
XX
XX This invention relates to a novel method for detecting a hereditary risk
CC for a coronary spasm. Specifically, it refers to the identification of
CC two or more polymorphisms and determining the genotype of the nucleic
CC acid sample in order to calculate the risk of the patient to a coronary
CC spasm. The present invention describes a risk analysis that comprises
CC identifying two or more polymorphisms occurring at any position taken
CC from -863 of the tumor necrosis factor alpha gene, 242 of the NADH/NADPH
CC oxidase p22 phox gene, -219 of the apolipoprotein E gene, -482 of the
CC apolipoprotein C-II gene, -1171 of the streptolysin 1 gene, 4070 of the
CC apolipoprotein E gene, -634 of the interleukin 6 gene, -786 of the
CC endothelium type nitrogen monoxide synthetase gene or position 2445 of
CC the fatty acid binding protein 2 gene. Accordingly, the method enables
CC prevention of cardiac sudden death resulting from variant angina
CC pectoris, coronary spasm and myocardial infarction. This polymorphic
CC is the human fatty acid binding protein 2 gene Alu repeat region
CC containing a G2445A SNP in the sequence given in the invention. NOTE:
CC This sequence is referred to as Seqid 9 in the sequence listing, but
CC Seqid 8 also refers to a primer given in the disclosure of the
CC specification.
XX
SQ Sequence 5204 BP; 1770 A; 867 C; 836 G; 1731 T; 0 U; 0 Other;
Query Match 2.1%; Score 49.8; DB 12; Length 5204;
Best Local Similarity 59.2%; Pred. No. 0.26;
Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;
QY 1516 TGAGCTTTAGCCAGCAGCATCATCATGTAATGCTTCTGTAAGCCTTCATTAAT 1575
DB 971 TGAACCTTTAGCTTCCACATCACTACATGATGAGTTGGTTCAAGATTAATTAATTAAT 1030
QY 1576 TCTC-----TTTGCAAGCTCTGCTACTTACCAAGCTGCTTACAGACAGAAAG 1627
DB 1031 TCTGCCCAAGAGACAGACCTGAAATCTTACTGCTCTTAGAGGCTGACT-CAACTGAATCA 1089
QY 1628 TGGCATTTAACGGTACTTGGAAAAATGAGAAAAATGAAACTTGTAAAAATTCATGGAAG 1687
DB 1090 TGGCGTTTACACGACTTGAAGGTAGACCGAGGTGAAAACTATGCAAGTTTCATGGAAA 1149
QY 1688 CAATGGTAAAGCCTTATTTTGAATGCT 1718
DB 1150 AATGGGTAAAGACTTATTCTTTGTGCT 1180
RESULT 7
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
XX ADA71938;
AC
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; de.
XX
XX Oryza sativa.
OS
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX

CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine IMF D ESR (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139
XX
SQ Sequence 337 BP; 150 A; 35 C; 40 G; 112 T; 0 U; 0 Other;
Query Match 2.0%; Score 47.2; DB 8; Length 337;
Best Local Similarity 53.9%; Pred. No. 0.44;
Matches 97; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 2185 AAGAGTAAGCTGCTGAGTCAATTCGAGCTACGCTCAACATACAGAAAGTCA 2244
DB 289 AATTAATATATATATATATTAACAATTAATAATTAAGTATATATATAAAGATGAT 230
QY 2245 AATTAATAATGCAACATGATGTTCTATTTTCTGCTGATGTTAAATGATAT 2304
DB 229 ATTAATATATATATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTT 170
QY 2305 TATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2364
DB 169 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 110
RESULT 9
ABLI5880
ID ABLI5880 standard; cDNA; 2636 BP.
XX
XX ABLI5880;
AC
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 42122.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR P-PDB; ABB71777.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 42122; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABJ30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 2636 BP; 849 A; 414 C; 497 G; 876 T; 0 U; 0 Other;
Query Match 2.0%; Score 47; DB 4; Length 2636;
Best Local Similarity 52.3%; Pred. No. 0.97;
Matches 104; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 1821 ATTGCCCTGCACTTATTTTGAAGAAAGCTATAGAGGGAATACAAAGAAACAT 1880
DB 540 ATGACCTTGAATATCTTAAATATACAGCCCTGTTCTGTTCTTCAATTTAAACAT 599
QY 1881 CTGATTTTATTTGCAATTCGATATCTTATGCAATTTAGCTAATTCAGTAGAGCATTC 1940
DB 600 ATTAATTTATTTAAATGGAATTAATTTGTTCTTAAATAATTCAGGATATTTATTTTC 659
QY 1941 ACCAGAAATTTAATAGATTTATATGTAAGCAATTTATTTGTAATGACTTTGAAA 2000
DB 660 TTAGGGAATTAAGATTAATTTGCTGCTAGTTTATATGTTAAAAAAGGTATAGACA 719
QY 2001 ATTACACAAGAGGAAT 2019
DB 720 ATTAGTAAGTGTATATTT 738
RESULT 10
ADD29587
ID ADD29587 standard; mRNA; 662 BP.
XX
XX ADD29587;
AC
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse tumour suppressor mRNA SEQ ID NO:36.
XX
XX ss; mouse; tumour suppressor; cancer; cancer; cytostatic; gene therapy.
XX
OS Mus musculus.
XX
PN WO2003058201-A2.
XX
PD 17-JUL-2003.
XX
PF 31-DEC-2002; 2002WO-US041825.
XX
PR 31-DEC-2001; 2001US-0345317P.
XX
PA (QUAR-) QUARK BIOTECH INC.
PA (CLEV-) CLEVELAND CLINIC FOUND.
XX
PI Feinstein E, Gudkov AV;
XX
DR WPI; 2003-598393/56.
XX
PT Diagnosing cancer comprises determining the polypeptide or polynucleotide
PT levels e.g., hepatic lipase, in a sample from a subject, where a higher
PT level compared to that in a subject free of cancer is indicative of
PT cancer.
XX
PS Disclosure; SEQ ID NO 36; 272pp; English.
XX
XX The invention relates to a novel method for diagnosing a cancer in a
CC subject. the method comprises determining, in a sample from the subject,
CC the level of at least one polypeptide, where a higher level of the
CC polypeptide compared to the level of the polypeptide in a subject free of
CC cancer is indicative of cancer. The polypeptide is selected from any of
CC the polypeptides encoded by the polynucleotides listed in the
CC specification and polypeptides which are at least 70% homologous to the
CC polypeptides. The method of the invention has cytostatic activity, and
CC may have a use in gene therapy. The method is useful in identifying
CC markers specific for one or several types of cancer, depending on the
CC tissue origin, which may be used in numerous diagnostic and prognostic
CC applications as well as cancer type-specific targets for therapeutic
CC intervention. The compounds that modulate the activity of a tumour

CC suppressor gene are useful in the treatment of cancer or as anti-cancer
CC drugs. The present sequence represents a polynucleotide of the invention.
XX
SQ Sequence 662 BP; 217 A; 133 C; 148 G; 164 T; 0 U; 0 Other;
Query Match 2.0%; Score 46.8; DB 10; Length 662;
Best Local Similarity 73.2%; Pred. No. 0.68; Mismatches 22; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 1615 ACAGACAGAAAGATGCGATTAAACGTAATTGAGAAATAGAGAAATGAGAACTATGAA 1674
DB 28 ACAGCTGAGATCATGCGATTGACGCGACGTGGAAGTAGACCGGAAACGAGAACTATGAA 87
QY 1675 AATTCATGAGACGATCGGTA 1696
DB 88 AAGTTCATGAGAGAAATGCGCA 109
RESULT 11
ID AAL19953 standard; cDNA; 610 BP.
XX AAL19953;
XX AAL19953;
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 12410.
XX
XX Human breast cancer; cell marker; cytostatic; ss.
XX Homo sapiens.
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
XX 14-MAR-2000; 2000US-0189167P.
XX 24-MAR-2000; 2000US-0192099P.
XX 29-MAR-2000; 2000US-0193480P.
XX 15-MAY-2000; 2000US-0205230P.
XX 09-JUN-2000; 2000US-0211315P.
XX 25-JUL-2000; 2000US-0220534P.
XX
XX (MIL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX MPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX Claim 1; Page 2193; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AAL07544-AAL26789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterizing treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cytostatic
XX activity
XX
SQ Sequence 610 BP; 211 A; 93 C; 96 G; 205 T; 0 U; 5 Other;
Query Match 2.0%; Score 46.6; DB 4; Length 610;
Best Local Similarity 62.4%; Pred. No. 0.74; Mismatches 44; Indels 0; Gaps 0;
Matches 73; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 2204 AGAATTCGGGCTAAGCTCCCTCAACTACAGAAAAAGTCATTAATAATGCAAACTGA 2263

DB 267 AAAAGGCTGTTTACCTTTTTCATGCGGAAAAAAGAAAAAAGAAAAA 208
QY 2264 TGTTCATTTTGTGTTTCTCGCTGATGTTAATGATTATTAATTTTTTTT 2320
DB 207 AGGCGCTTT 151
RESULT 12
ID ABL70479 standard; DNA; 9742 BP.
XX ABL70479;
XX ABL70479;
XX 01-JUL-2002 (first entry)
XX
XX Chemically treated cell signalling DNA sequence#185.
XX
XX Cell signalling; cytosine methylation; cell signalling disease; cancer;
XX tumour; cytostatic; ds.
XX
XX Unidentified.
XX
XX WO200202807-A2.
XX
XX 10-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-EP007471.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX MPI; 2002-154758/20.
XX
XX Nucleic acid, useful for diagnosis and therapy of diseases associated
XX with cell signalling e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with cell signalling.
XX
XX Claim 1; SEQ ID NO 369; 24pp + Sequence listing; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of at least
XX 18 bases of a segment of chemically pretreated DNA of genes associated
XX with cell signalling. The activity of the modified sequences of the
XX invention may be described as cytostatic. The object of the invention is
XX to provide the chemically modified DNA of genes associated with cell
XX signalling, as well as oligonucleotides and/or PNA-oligomers for
XX detecting cytosine methylations, as well as a method which is
XX particularly suitable for the diagnosis and/or therapy of genetic and
XX epigenetic parameters of genes associated with cell signalling. The
XX chemically modified DNA provided by the invention is useful for diagnosis
XX and therapy of diseases such as solid tumours and cancer. The sequences
XX given in records ABL70111-ABL70626 represent chemically pre-treated
XX genomic DNA's of genes associated with cell signalling. Note: The
XX sequence data for this patent is not represented in the printed
XX specification, but is based on sequence information supplied by the
XX European Patent Office
XX
SQ Sequence 9742 BP; 3039 A; 88 C; 1939 G; 4676 T; 0 U; 0 Other;
Query Match 2.0%; Score 46.6; DB 6; Length 9742;
Best Local Similarity 47.7%; Pred. No. 1.9; Mismatches 149; Indels 0; Gaps 0;
Matches 136; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 2051 TTGAGTCATTTTAACATGCTAGTCTTAAGCTTGTACAAACGCTGTAAATATG 2110
DB 6826 TTAGTATTTTGTGTAAGATATATTTTAAATTTTATTTTGTATATATATA 6885
QY 2111 GATGTAAATTAATCTCTAGTGTAGATAGTGTGTTTATACAGGCTAACACCTGCCTC 2170

```
D6      TATGTAGATTTTTAAATAAATTGANAAGAAGTGTAATTATTAACAGTAGATTTTTTA    6948
Q7      AGCGAAAAGTCGCAGACAAGTAAGACTCTGACTCAGAATTCTGGCCTAACGCTCCCTACT    2230
D8      GGTATATGTTGGTAATTAATTAATTGTTGTTTTTATTTTAATTATGATATTTTTTATTT    7005
Q9      ACACAAAAAGCACAGATPAAAAATGCCAACAATGATGCTCATTTTGTCTTCTCGCTGA    2290
D9      ATTTTATTTATTTAAGATATATTTGTTTACGTGAATGTTATGTTTTTTTTTTTTTTTTTT    7065
QY      2291   TGTTAATTGATTATTAATTTATTTTTTTTTTTTAAAGCGTAATGTGAT   2335
Db       |---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
          7066 TATTATGTTGTTGTAATTTTTTTGTTTTTTTATTTTATTAAGGCT   7110

RESULT_13
ABLJ32692 ID ABLJ32692 standard; DNA; 6222 BP.
XX ABJ32692;
AC     ABLJ32692;
DT     26-MAR-2002 (first entry)
DE     Human immune system associated gene SEQ ID NO: 665.
XX     XX
KM     Human, immune system disease; cytosine methylation; antiasthmatic;
KW     antidiabeticosclerotic; antianemic; cytostatic; nootropic;
KW     neuroprotective; anti-HIV; anticovulsant; ophthalmological;
KW     antirheumatic; antiarthritis; antidiabetic; antiporiastic;
KW     antiflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KM     acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KM     neutofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
OS     ds.
XX     Homo sapiens.
XX OS WO200200928-A2.
PN PD 03-JAN-2002.
PF PF 02-JUL-2001; 2001WO-EP007537.
PR PR 30-JUN-2000; 2000DE-01032529.
PA PA 01-SEP-2000; 2000DE-01043826.
PI PI (EPIG-) EPIGENOMICS AG.
DR DR Olek A, Piepenbrock C, Berlin K;
SQ SQ WPI; 2002-130909/17.

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukemia, Alzheimer's disease, AIDS, epilepsy, neutrofibrmatosis, CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention
CC CC
SQ Sequence 6222 BP; 1941 A; 55 C; 1168 G; 3058 T; 0 U; 0 Other;
```

[illegible]

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2005, 19:03:25 ; Search time 6960 Seconds

(without alignments)
13021.699 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381

Sequence: 1 agctctctgcgcagaaag.....atctgaagctcactatcag 2381

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	752.6	31.6	1473	8	CC250935	CH261-17C
2	99	4.2	598	7	CF250818	esa017_eo
3	95	4.0	595	7	CF250947	esa019_a0
4	85	3.6	705	5	BU355106	BU355106 603474288
5	84	3.5	759	5	BU265304	BU265304 603508640
6	81	3.4	829	5	BU296697	BU296697 603741965
7	79.4	3.3	885	5	BU123280	BU123280 603003274
8	79	3.3	814	5	BU234099	BU234099 603792609
9	78.8	3.3	1080	5	BU123336	BU123336 603149475
10	75.8	3.2	426	6	CD739695	CD739695 4028031.1
11	74	3.1	576	6	CD739141	CD739141 4026481.1
12	70.4	3.0	480	6	CD733346	CD733346 4045132.1
C 13	62.4	2.6	1101	9	CNS00395	AL063921 Drosophila
14	61.4	2.6	885	5	BU123717	BU123717 603147391
15	60.4	2.5	443	5	BK276255	BK276255 BK276255
16	57.2	2.4	683	7	CN793516	CN793516 4128553.1
17	55.8	2.3	657	7	CK945920	CK945920 4070423.1
18	54.2	2.3	572	7	CK834687	CK834687 4059069.1
19	54.2	2.3	673	7	CK947902	CK947902 4072726.1
20	54.2	2.3	686	7	CK957350	CK957350 4097834.1
21	54.2	2.3	708	7	CK960577	CK960577 4101917.1
22	54	2.3	481	4	BM430667	BM430667 IDU03A11
23	54	2.3	489	6	CB224035	CB224035 IDU07F03
24	54	2.3	497	4	BM430956	BM430956 IDU07F03

25	54	2.3	548	4	BM432416	BM432416 IJEU4B9.a
26	54	2.3	568	4	BM431104	BM431104 IDU09G09
27	54	2.3	671	7	CK980160	CK980160 4111791.1
28	54	2.3	679	7	CK947198	CK947198 4071820.1
29	54	2.3	682	7	CK948900	CK948900 4074042.1
C 30	54	2.3	684	7	CK947048	CK947048 4071436.1
31	54	2.3	702	7	CK944155	CK944155 4068182.1
32	54	2.3	720	7	CK946512	CK946512 4071131.1
33	54	2.2	372	4	BM430450	BM430450 IDU034E2
34	53	2.2	414	4	BM433066	BM433066 IJEUJ1H7
35	53	2.2	416	4	BM432621	BM432621 IJEUJ1H7
36	53	2.2	546	7	CN098758	CN098758 EC2CA15A
C 37	53	2.2	576	7	CN098757	CN098757 EC2CA15A
38	52.6	2.2	660	7	CK954372	CK954372 4094190.1
39	52.6	2.2	1101	9	CNS00395	AL063921 Drosophila
40	52.4	2.2	503	4	BM432863	BM432863 IJEUJ1H7
41	52	2.2	339	4	BM430492	BM430492 IDU035B12
42	52	2.2	371	4	BM432453	BM432453 IJEUJ1H7
43	52	2.2	371	4	BM433013	BM433013 IJEUJ1H7
44	52	2.2	415	4	BM430284	BM430284 IDU032C9
45	52	2.2	453	4	BM431869	BM431869 IJEUJ1H7

ALIGNMENTS

RESULT 1
LOCUS CC250935 1473 bp. DNA linear GSS 13-MAY-2003
DEFINITION CH261-17C7 Sp6.1 CH261 Gallus gallus genomic clone CH261-17C7,
genomic survey sequence.
ACCESSION CC250935
VERSION CC250935.1 GI.30587685
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 1473)
AUTHORS Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 18200 Std Error: 0.00
Class: BAC ends
Seq primer: Sp6 ATTGAGTGACACTATAG
High quality sequence start: 239
High quality sequence stop: 912.
Location/Qualifiers
1. 1473
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-17C7"
/sex="female"
/cell_line="VUCD001, inbred 256"
/note="Vector: PTABAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - For library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match 31.6%; Score 752.6; DB 8; Length 1473;
Best Local Similarity 98.3%; Pred. No. 4.3e-168;
Matches 793; Conservative 0; Mismatches 9; Indels 5; Gaps 3;
QY 1580 TTTCGAAGCTGCTACTTACCAAGATC-TGCCACAGACAGAAAGATGCATTAA-1637

```

Db      1042 TTGCAGAAAGTTGCTCTTCCAGAGAAATCTCCCTACAGACAGAAAGATGGCATTTTAAAC 983
      1638 CCGTACTTGAAAATATGAGAAAAATGAGAACTATGAAAAATTCATGGAAGCAATGGGTAA 1697
      982 CCGTACTTGAAAATATGAGAAAAATGAGAACTATGAAAAATTCATGGAAGCAATGGGTAA 923
      1698 GCCTTACTTTTGAATGCTTCTTAAGACAGATACCACTACGGGGGAATACAACTTA 1757
      922 GCCTTACTTTTGAATGCTTCTTAAGACAGATACCACTACGGGGGAATACAACTTA 863
      1758 AGCTGTTCAGAACTACCACTCTGCTACCTGCTCTTGTGCTGCTATTTTGGCCCTTG 1817
      862 AGCTGTTCAGAACTACCACTCTGCTACCTGCTCTTGTGCTGCTATTTTGGCCCTTG 803
      1818 CACATTCCTCTGCACTTATTTTGAAGAACTCTATAGAGGGGAATACAAAGAAAGAAAC 1877
      802 CACATTCCTCTGCACTTATTTTGAAGAACTCTATAGAGGGGAATACAAAGAAAGAAAC 743
      1878 ATTCTGATTTTATTTGCAATTCGATATCTTATGCAATTTAGCTAATTCAGATGAGGCAT 1937
      742 ATTCTGATTTTATTTGCAATTCGATATCTTATGCAATTTAGCTAATTCAGATGAGGCAT 683
      1938 TCCAGAGAAATTTAATAGAAATTTATGTAAGAAATTTATTTGATTAAGACTGTTTGA 1997
      682 TCCAGAGAAATTTAATAGAAATTTATGTAAGAAATTTATTTGATTAAGACTGTTTGA 623
      1998 AAAATTTACACAGAGAGAAATTTGCTGCTCTCCAGTTTTCAGAACACACATGTTGAGT 2057
      622 AAAATTTACACAGAGAGAAATTTGCTGCTCTCCAGTTTTCAGAACACACATGTTGAGT 563
      2058 CATTTAACATGCTAGTCTTACTTATAGCTGTGTACAACTGCTGTATATGATGATGTA 2117
      562 CATTTAACATGCTAGTCTTACTTATAGCTGTGTGTACAACTGCTGTATATGATGATGTA 503
      2118 ACATACTATCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2177
      502 ACATACTATCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443
      2178 GGTGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2237
      442 GGTGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 383
      2238 AAGTCAATTAATAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2297
      382 AAGTCAATTAATAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 323
      2298 TGATTTATTTA---TTTTTTTTTTTGAAGGCTGATGATGATGATGATGATGATGATGATGAT 2354
      322 TGATTTATTTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 263
      2355 CACGATATCTGAGCTCACTATTCAG 2381
      262 CACGATATCTGAGCTCACTATTCAG 236

RESULT 2
LOCUS   CF250818      598 bp      mRNA      linear      EST 07-AUG-2003
DEFINITION  esao17_e05 Elmeiria tenella-infected caecal tonsil Gallus gallus
ACCESSION  CF250818
VERSION    CF250818.1 GI:33484073
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 598)
            Witzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J.,
            Chausse,A.M. and Zoorob,R.
            A collection of chicken ESTs from activated immune cells
  
```

```

JOURNAL   Unpublished (2003)
COMMENT    Contact: Zoorob R
            UPR 1983
            CNRS
            7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France
            Tel: 33 1 49 58 35 00
            Fax: 33 1 49 58 33 81
            Email: zoorob@vjf.cnrs.fr.
FEATURES   Location/Qualifiers
            source          1..598
                        /organism="Gallus gallus"
                        /mol_type="mRNA"
                        /db_xref="taxon:9031"
                        /clone_lib="Elmeiria tenella-infected caecal tonsil"
                        /note="Organ: Caecal tonsil; Vector: pTribEx2"

ORIGIN
Query Match      4.2%; Score 99; DB 7; Length 598;
Best Local Similarity 100.0%; Pred. No. 1,9e-12;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      QY      1596 ACTTACCAAGTCTGCTTACAGACAGAAAGATGCGATTTAAGCTTGGAAAAATAGA 1655
      DB      3 ACTTACCAAGTCTGCTTACAGACAGAAAGATGCGATTTAAGCTTGGAAAAATAGA 62
      QY      1656 GAAAAATGAGACTATGAAAAATTCATGAGAACCAATGG 1694
      DB      63 GAAAAATGAGACTATGAAAAATTCATGAGAACCAATGG 101

RESULT 3
LOCUS   CF250947      595 bp      mRNA      linear      EST 07-AUG-2003
DEFINITION  esao19_a07 Elmeiria tenella-infected caecal tonsil Gallus gallus
ACCESSION  CF250947
VERSION    CF250947.1 GI:33484202
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 595)
            Witzell,H., Bed'Hom,B., Morin,V., Young,J.R., Whitaker,C.J.,
            Chausse,A.M. and Zoorob,R.
            A collection of chicken ESTs from activated immune cells
            JOURNAL   Unpublished (2003)
            COMMENT    Contact: Zoorob R
            UPR 1983
            CNRS
            7 rue Guy Moquet, BP 8, 94801 Villejuif cedex, France
            Tel: 33 1 49 58 35 00
            Fax: 33 1 49 58 33 81
            Email: zoorob@vjf.cnrs.fr.
FEATURES   Location/Qualifiers
            source          1..595
                        /organism="Gallus gallus"
                        /mol_type="mRNA"
                        /db_xref="taxon:9031"
                        /clone_lib="Elmeiria tenella-infected caecal tonsil"
                        /note="Organ: Caecal tonsil; Vector: pTribEx2"

ORIGIN
Query Match      4.0%; Score 95; DB 7; Length 595;
Best Local Similarity 100.0%; Pred. No. 1,7e-11;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      QY      1600 ACCAGAACTCTGCTTACAGACAGAAAGATGCGATTTAAGCTTGGAAAAATAGA 1659
      DB      3 ACCAGAACTCTGCTTACAGACAGAAAGATGCGATTTAAGCTTGGAAAAATAGA 62
      QY      1660 AATGAGAACTATGAAAAATTCATGAGAACCAATGG 1694
  
```

Db 63 AATGAGACTATGAAAAATTCAATGAGCAATGG 97

RESULT 4
LOCUS BU2653106
DEFINITION 603474288F1 CSEQCHN70 Gallus gallus cDNA clone CHEST355022 5', mRNA
SEQUENCE
ACCESSION BU2653106
VERSION BU2653106.1 GI:25863107
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J. A Comprehensive Collection of Chicken cDNAs Curr. Biol. 12 (22), 1965-1969 (2002)
JOURNAL MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..705
location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST355022"
/dev_stage="36"
/lab_host="DH10B"
/clone_1b="CSEQCHN70"
/note="Organ: hearts; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
Query Match 3.6%; Score 85; DB 5; Length 705;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 TGCCTACAGACAGAAAGATGGCATTTAAGCTTGAAGAAATAGAGAACT 1669
DB 1 TGCCTACAGACAGAAAGATGGCATTTAAGCTTGAAGAAATAGAGAACT 60
QY 1670 ATGAAAAATTCATGAGCAATGGG 1694
DB 61 ATGAAAAATTCATGAGCAATGGG 85

RESULT 5
BU265304

LOCUS BU265304 759 bp mRNA linear EST 26-NOV-2002
DEFINITION 603508640F1 CSEQCHN52 Gallus gallus cDNA clone CHEST437010 5', mRNA
SEQUENCE
ACCESSION BU265304
VERSION BU265304.1 GI:25536254
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J. A Comprehensive Collection of Chicken cDNAs Curr. Biol. 12 (22), 1965-1969 (2002)
JOURNAL MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..759
location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST437010"
/dev_stage="22"
/lab_host="DH10B"
/clone_1b="CSEQCHN52"
/note="Organ: limbs; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
Query Match 3.5%; Score 84; DB 5; Length 759;
Best Local Similarity 100.0%; Pred. No. 7.4e-09;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1611 GCTTACAGACAGAAAGATGGCATTTAAGCTTGAAGAAATAGAGAACTA 1670
DB 1 GCTTACAGACAGAAAGATGGCATTTAAGCTTGAAGAAATAGAGAACTA 60
QY 1671 TGAATAATTCATGAGCAATGGG 1694
DB 61 TGAATAATTCATGAGCAATGGG 84

RESULT 6
BU296697

LOCUS BU296697 829 bp mRNA linear EST 27-NOV-2002
DEFINITION 603741965F1 CSEQCHN56 Gallus gallus cDNA clone CHEST4017 5', mRNA
SEQUENCE
ACCESSION BU296697
VERSION BU296697.1 GI:25746333
KEYWORDS EST.

Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

Location/Qualifiers

1..814

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, H1sex"

/db_xref="taxon:9031"

/clone="CHEST758b4"

/dev_stage="22"

/lab_host="DH10B"

/clone_1ib="CSBOCHN24"

/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:

EcORI; Site_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methyated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with

EcORI, size-selected, and cloned into the NotI and EcORI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldi et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

ORIGIN

Query Match 3.3%; Score 79; DB 5; Length 814;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1616 CAGACGAGAAAGATGCGATTAAAGCGTACTGGAATAAGAAATGAGAACTATGAAA 1675
|||||
DB 1 CAGACGAGAAAGATGCGATTAAAGCGTACTGGAATAAGAAATGAGAACTATGAAA 60
|||||

QY 1676 AATTCAATGAGCAATGCG 1694
|||||
DB 61 AATTCAATGAGCAATGCG 79
|||||

RESULT 9

BU123336

LOCUS BU123336 1080 bp mRNA linear EST 25-NOV-2002

DEFINITION 603189475F1 CSBQCHL18 Gallus gallus cDNA clone CHEST15319 5', mRNA

sequence.

ACCESSION BU123336

VERSION BU123336.1

KEYWORDS GI:25333972

SOURCE EST.

ORGANISM Gallus gallus (chicken)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 1080)

Boardman, P.E., Sanz-Eguerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A., and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

1..1080

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Compton line 151"

/db_xref="taxon:9031"

/clone="CHEST15319"

/sex="Female"

/dev_stage="adult"

/lab_host="DH10B"

/clone_1ib="CSBQCHL18"

/note="Organ: small intestine; Vector: pBluescript II

KS(+); Site_1: EcORI; Site_2: NotI; Modification of

pBluescript II KS(+). (Stratagene) vector to accommodate

cDNA produced with the T-primed protocol (Construction of

uni-directionally cloned cDNA libraries from messenger RNA

for improved 3' end DNA sequencing by Glenn Py, et al.

U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with

NotI and EcORI. Ligate in double stranded adaptor

containing BspI and BamHI sites

[5'ggcgcgtcagcccgatccgaaataag]

[5'aattctttttcggatccggtcgacgc]"

ORIGIN

Query Match 3.3%; Score 78.8; DB 5; Length 1080;
Best Local Similarity 97.6%; Pred. No. 1.4e-07;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1613 CTACAGACGAAAGATGCGATTAAAGCGTACTGGAATAAGAAATGAGAACTATG 1672
|||||
DB 7 CAGACGAGAAAGATGCGATTAAAGCGTACTGGAATAAGAAATGAGAACTATG 66
|||||

QY 1673 AAAAATTCATGAGCAATGCG 1694
|||||
DB 67 AAAAATTCATGAGCAATGCG 88
|||||

RESULT 10

CD739695

LOCUS CD739695 426 bp mRNA linear EST 26-JUN-2003

DEFINITION 4028031 1GAL - Chicken Intestinal Lymphocyte Gallus gallus cDNA

clone 1GAL_20B02 5', mRNA sequence.

ACCESSION CD739695

VERSION CD739695.1

KEYWORDS GI:32290544

SOURCE EST.

ORGANISM Gallus gallus (chicken)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 426)

Min, W., Liliehoj, H.S., Ashwell, C.M., Matukumalli, L.K., van

Tassel, C. and Han, J.Y.

Chicken Intestinal Lymphocyte EST database as a resource for the

analysis of mucosal immune function

Unpublished (2003)

Contact: Hyun S. Liliehoj

Animal Parasite Diseases Laboratory

Animal and Natural Resources Institute, USDA

BLDG.1043, BARC-East, Beltsville, MD 20705, USA

Tel: 3015048771

Fax: 3015045103

Email: hlillehoj@anri.barc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

0.000925 using options -trim_alt '' -trim_fasta. Vector identified

by cross_match using options -mismatch 12 -mnscore 12

Plate: 20 row: B column: 02

Seq primer: ATTTAGGTGACACTATAG

High quality sequence stop: 426.

Location/Qualifiers

1..426

/organism="Gallus gallus"

/mol_type="mrna"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL_20802"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/lab_host="EMDH10B"
/note="Organ: Intestine; Vector: PCMV-SPORT6; Site: 1; SalI; Site 2: NotI; Normalized library from chicken gut infected with coccidia duodenum and middle gut."
ORIGIN

Query Match 3.2%; Score 75.8; DB 6; Length 426;
Best Local Similarity 97.5%; Pred. No. 5.8e-07;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1616 CAGACGAAAGATGGCATTTAAGCGTACTTGAAGAAATAGAGAAATGAGAACTATGAAA 1675
|||||
DB 2 CAGACGAAAGATGGCGTTTAAACGCTTCTTGAAATAGAGAAATGAGAACTATGAAA 61
|||||

QY 1676 AATTGATGGAAGCAATGGG 1694
|||||
DB 62 AATTGATGGAAGCAATGGG 80
|||||

RESULT 11
CD7339141 576 bp mRNA linear EST 26-JUN-2003
LOCUS 4026481 IGAL - Chicken Intestinal Lymphocyte Gallus gallus cDNA
DEFINITION
ACCESSION CD7339141 GI:32289990
VERSION CD7339141.1 GI:32289990
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 576)
Min.W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van
Tassel,C. and Han,J.Y.
Chicken intestinal lymphocyte EST database as a resource for the
analysis of mucosal immune function
Unpublished (2003)
CONTACT: Hyun S. Lillehoj
Animal Parasite Diseases Laboratory
Animal and Natural Resources Institute, USDA
Bldg. 1043, BARC-East, Beltsville, MD 20705, USA
Tel: 3015048771
Fax: 3015045103
Email: hlilleho@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim_alt '' -trim_fasta. Vector identified
by cross_match using options -mismatch 12 -minscore 12
Plate: 93 row: A column: 12
Seq primer: ATTTAGGTGACACTATAG
High quality sequence stop: 576.
Location/Qualifiers
1. 576
/organism="Gallus gallus"
/mol_type="mrna"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL_93A12"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_1tb="IGAL - Chicken Intestinal Lymphocyte"

FEATURES
source

/note="Organ: Intestine; Vector: PCMV-SPORT6; Site: 1; SalI; Site 2: NotI; Normalized library from chicken gut infected with coccidia duodenum and middle gut."
ORIGIN

Query Match 3.1%; Score 74; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1621 AGAAGATGGCATTTAAGCGTACTTGAAGAAATAGAGAAATGAGAACTATGAAAATTC 1680
|||||
DB 1 AGAAGATGGCATTTAAGCGTACTTGAAGAAATAGAGAAATGAGAACTATGAAAATTC 60
|||||

QY 1681 ATGGAAGCAATGGG 1694
|||||
DB 61 ATGGAAGCAATGGG 74
|||||

RESULT 12
CD733346 480 bp mRNA linear EST 26-JUN-2003
LOCUS 4045132 IGAL - Chicken Intestinal Lymphocyte Gallus gallus cDNA
DEFINITION
ACCESSION CD733346
VERSION CD733346.1 GI:32284195
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 480)
Min.W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van
Tassel,C. and Han,J.Y.
Chicken intestinal lymphocyte EST database as a resource for the
analysis of mucosal immune function
Unpublished (2003)
CONTACT: Hyun S. Lillehoj
Animal Parasite Diseases Laboratory
Animal and Natural Resources Institute, USDA
Bldg. 1043, BARC-East, Beltsville, MD 20705, USA
Tel: 3015048771
Fax: 3015045103
Email: hlilleho@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim_alt '' -trim_fasta. Vector identified
by cross_match using options -mismatch 12 -minscore 12
Plate: 41 row: E column: 15
Seq primer: ATTTAGGTGACACTATAG
High quality sequence stop: 480.
Location/Qualifiers
1. 480
/organism="Gallus gallus"
/mol_type="mrna"
/strain="white leghorn SC"
/db_xref="taxon:9031"
/clone="IGAL_41B15"
/sex="mixed"
/tissue_type="Gut"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_1tb="IGAL - Chicken Intestinal Lymphocyte"
/note="Organ: Intestine; Vector: PCMV-SPORT6; Site: 1; SalI; Site 2: NotI; Normalized library from chicken gut infected with coccidia duodenum and middle gut."
ORIGIN

Query Match 3.0%; Score 70.4; DB 6; Length 480;
Best Local Similarity 98.6%; Pred. No. 1.2e-05;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1623 AAGATGCGATTTAAGCGTACTTGAAGAAATAGAGAAATGAGAACTATGAAAATTCAT 1682

DB	LOCUS	DEFINITION	ACCESSION	VERSION	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	ORIGIN
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1999							
Db	1683	GGAGCAATGGG 1694											
Db	61	GGAGCAATGGG 72											
Db	1	AAAGTCGCAATTTAAACGGCAGCTTGAGAAATAGAGAAATGAGAACTATGAAAAATTCAT	1101 bp	DNA	linear	GSS 03-JUN-1							

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
2082	TTAAGCTTTGTAACAACACTGCTCGTAAATATGATGTAAACATPACTACTGTTGGAAT	2141				
745	DDRRAGTAGAGKMMRRTRKRRKRRKRDFTWDDADADDTARDDRRRGGDGAAGKGTGRK	686				
2142	AGTTTGATTTTCAAGCGCTGAACACTGCGCTCAGGTGAAGAGTGAAGAGTAAGCTCGAG	2201				
685	RRRDRATWDTTDWADWADAAWTTTDTDTDDWGRDRRRKRRKRRRTTAAADWWTWK	626				
2202	TCAGAAATTCGGGCTAAAGCTCCCTCAACTACAGAAAAAGTCACATAAAATGCAACAT	2261				
625	-AMDMAKMDKTRADBRMDRWADATWTDARADRWAAABARABARABARABARADRRWTK	567				
2262	GATGTTCTATTGTTGTTTCTCTGCTGGAATTAATGATTATTAATTTTATTTT	2321				
566	GRTTATATWTTWAAATAAAWAAWATTTATATWTTTWTWTTTTTTTTTTTAAWMAA	507				
2322	GGCGTGAATGATGATGAATAAGAAATTT	2347				
506	WWTATWAAWTTAAWAAAAAATAATTT	481				
RESULT 14						
LOCUS	BUI23717	885 bp	mRNA	linear	EST 25-NOV-2002	
DEFINITION	601247391.F1 CSEQCHL18 Gallus gallus CDNA clone CHEST146124 5', mRNA					
ACCESSION	BUI23717					
VERSION	BUI23717.1	GI:25334366				
KEYWORDS	EST.					
SOURCE	Gallus gallus (chicken)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.					
	1 (bases 1 to 885)					
	Boardman, P.E., Sans-Ezquerro, J., Overton, I.M., Burt D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.					
	A Comprehensive Collection of Chicken CDNAs					
	Curr. Biol. 12 (22), 1965-1969 (2002)					
	12445392					
	Contact: Simon Hubbard					
	Department of Biomolecular Sciences					
	University of Manchester Institute of Science and Technology					
	(UMIST)					
	PO Box 88, Manchester, M60 1QD, UK					
	Tel: 01612008930					
	Fax: 01612360409					
	Email: Simon.Hubbard@umist.ac.uk.					
FEATURES						
Source	Location/Qualifiers					
	1..885					
	/organism="Gallus gallus"					
	/mol_type="mRNA"					
	/strain="Compton Line 151"					
	/db_xref="taxon:9031"					
	/clone="CHEST146124"					
	/sex="Female"					
	/dev_stage="adult"					
	/lab_host="DH10B"					
	/clone_lib="CSEQCHL18"					
	/note="Organ: small intestine; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; Modification of pBluescript II KS(+): [Stratagene] vector to accommodate cDNA produced with the T-tipped protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BglI and BamHI sites					
	[5'ggcgcgtagccgcgcggatccgcaaaaag]					
	[5'aattcttttttcgatacggggtcgagcgc]"					

Query Match	2.6%	Score 61.4;	DB 5,	Length 885;
Best Local Similarity	96.1%;	Pred. No.	0.0019;	
Matches	74;	Conservative	0;	Mismatches 1; Indels 2; Gaps 1;
Oy	1618	GACAGAAAGATGGCTTTAACGCGTACTTGGAATAAGAGAAAAATGAGAACAATTGAAAAA	1677	
Dd	8	GAGAGAAAAGATGGCCTTTAACGCGTACTTGAAAAAT--AGAAAAATGAGAACTATGAAAAA	65	
Oy	1678	TTCATGGAAGCAATGGG	1694	
Dd	66	TTCATGGAAGCAATGGG	82	

```

Db      31 TGGGCGGTAATGTGATGTGAAAAGAAAGTTAGAGACCCACAGTATATCTGAAGCTACTATTC 90
QY      2380 AG 2381
          ||
Db      91 AG 92

Search completed: November 21, 2005, 01:59:31
Job time : 6965 secs

```

RESULT 15				
BX276255				
LOCUS				
DEFINITION	BX276255	443 bp	mRNA	linear EST 15-JUL-2004
	BX276255	AGENAE	Gallus gallus multi-tissues	normalized library

ACCESSION	BX276255	GI:28598746
VERSION	BX276255.1	
KEYWORDS	EST.	
SOURCE	Gallus gallus (chicken)	
ORGANISM	Gallus gallus	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 443)	Herault, F., le Meunh-Wetinger, V., Desser, C., Retout, E., Plum, F., Klop, C. and Douaire, M.	Construction and primary characterization of chicken normalized multi-tissue cDNA libraries	Unpublished (2003)	Contact: Douaire M

INRA, UMR INRA-ENSAR Genetique Animale
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
Tel: +33 (0) 2.23.48.54.63
Fax: +33 (0) 2.23.48.54.70
Email: Madeleine.Douaie@roazon.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at sigensupport@jouy.inra.fr to obtain the chromatogram of this
sequence.

row	c	column
24		

Plate: 0008
Seq primer: M13R.

FEATURES	Location/Qualifiers
source	1. .443

```

/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gcag0008c.c.24"
/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_id="AGE9NB Gallus gallus multi-tissues normalized library (gcag)
/notes="vector: pUT73D-pac; tissues: brain, embryos, kidney, multi-tissues, muscle, pancreas, skin, testis, liver, adipose tissue, granulosa, utero-vaginal gland, oviduct, small follicle, ovary, hypodochlamus, pituitary gland, ileon, jejunum, caecum, duodenum, spleen, proventriculus gland, bone marrow, thymus, hematopoietic progenitor cells. Clone distribution : AGE9NB Resource Centre, Francois PLOMI, Francois.Plomi.inra.fr, INRA, CEA, Radiobiologie et Etude du genome (LREB), Domaine de l'Ilvert, 78352, Jouy-en-Josas cedex, FRANCE"

```

ORIGIN

Query Match	2.5%	Score 60.4;	DB 5;	Length 443;
Best Local Similarity	98.4%	Pred. No. 0.0027;		
Matches 61; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy 2320 TAGGCGTGAATGTGATGAAAAAGAAAGTTAGAGGCCACGATATCTGAAAGCTCATATTC 2379

Search completed: November 21, 2005, 01:59:31
Job time : 6965 secs

Db 31 TGGCGGATGTGTGAAAAGAAAGTTAGAGCCACGATATCTGAAGCTCATTTC 90

QY 2380 AG 2381

Db 91 AG 92

Search completed: November 21, 2005, 01:59:31
Job time : 6965 secs

C	1	64.2	2.7	1141	4	US-09-806-7088-22	Sequence 22, Appl
C	2	57.8	2.4	1141	4	US-09-806-7088-22	Sequence 22, Appl
C	3	49.8	2.1	8905	4	US-09-949-016-11761	Sequence 11761, A
C	4	49.8	2.1	8907	4	US-09-949-016-16261	Sequence 16261, A
C	5	44.4	2.0	832	4	US-09-621-976-2813	Sequence 2813, A
C	6	48.2	1.9	515	3	US-09-385-982-474	Sequence 474, Appl
C	7	44.2	1.9	2252	4	US-09-949-016-4519	Sequence 4519, Appl
C	8	44.2	1.9	2273	4	US-09-949-016-19	Sequence 19, Appl
C	9	43.8	1.8	7218	1	US-08-232-463-14	Sequence 14, Appl
C	10	43	1.8	832	4	US-09-621-976-2813	Sequence 2813, A
C	11	42.6	1.8	601	4	US-09-949-016-18019	Sequence 18019, A
C	12	42.6	1.8	601	4	US-09-949-016-18020	Sequence 18020, A
C	13	42.6	1.8	601	4	US-09-949-016-161279	Sequence 161279, A
C	14	42.6	1.8	601	4	US-09-949-016-161280	Sequence 161280, A
C	15	42.6	1.8	611	3	US-09-385-982-376	Sequence 376, Appl
C	16	42.4	1.8	452	4	US-09-401-064-138	Sequence 138, Appl
C	17	42.2	1.8	652	4	US-09-270-767-10640	Sequence 10640, A
C	18	42.2	1.8	1257	4	US-09-270-767-15058	Sequence 15058, A
C	19	42	1.8	601	4	US-09-949-016-85025	Sequence 85025, A
C	20	42	1.8	601	4	US-09-949-016-85176	Sequence 85176, A
C	21	41.6	1.7	1273	4	US-09-270-767-14731	Sequence 14731, A
C	22	41.6	1.7	47363	4	US-09-949-016-173420	Sequence 13420, A
C	23	41.2	1.7	136	4	US-09-442-054-42	Sequence 42, Appl
C	24	41.2	1.7	136	4	US-09-442-054-42	Sequence 42, Appl
C	25	40.8	1.7	601	4	US-09-949-016-85023	Sequence 85023, A
C	26	40.8	1.7	601	4	US-09-949-016-85024	Sequence 85024, A
C	27	40.8	1.7	601	4	US-09-949-016-85026	Sequence 85026, A

C 28	40.8	1.7	601	4	US-09-949-016-85117	Sequence 85117, A
C 29	40.8	1.7	601	4	US-09-949-016-85175	Sequence 85175, A
C 30	40.8	1.7	601	4	US-09-949-016-85177	Sequence 85177, A
C 31	40.8	1.7	109159	4	US-09-949-016-14169	Sequence 14169, A
C 32	40.8	1.7	109159	4	US-09-949-016-14170	Sequence 14170, A
C 33	40.4	1.7	601	4	US-09-949-016-85027	Sequence 85027, A
C 34	40.4	1.7	601	4	US-09-949-016-85178	Sequence 85178, A
C 35	40.4	1.7	601	4	US-09-949-016-189998	Sequence 189998, A
C 36	40.4	1.7	192700	4	US-09-949-016-118920	Sequence 118920, A
C 37	40.4	1.7	192704	4	US-09-949-016-118920	Sequence 118920, A
C 38	40.4	1.7	640681	4	US-09-949-016-11182	Sequence 117182, A
C 39	40.2	1.7	1055	4	US-09-790-988-1	Sequence 1, Appl1
C 40	40.2	1.7	28862	4	US-09-806-708B-23	Sequence 23, Appl1
C 41	40.2	1.7	360470	4	US-09-949-016-16793	Sequence 16793, A
C 42	40	1.7	601	4	US-09-949-016-13171	Sequence 13173, A
C 43	40	1.7	601	4	US-09-949-016-146919	Sequence 146919, A
C 44	40	1.7	703	4	US-09-949-016-146920	Sequence 146920, A
C 45	40	1.7	733	4	US-09-945-016-146921	Sequence 146921, A
					US-09-270-767-10136	Sequence 10136, A

```

US-09-806-708B-22/c
: Sequence 22, Application US/09806708B
: Patent No. 6784342
:
: GENERAL INFORMATION:
: APPLICANT: The University of British Columbia
: TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
: FILE REFERENCE: 4810-58741
: CURRENT APPLICATION NUMBER: US/09/806,708B
: PRIORITY FILING DATE: 2001-04-03
: PRIOR APPLICATION NUMBER: US 60/147,133
: PRIORITY FILING DATE: 1999-08-04
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 22
:
: LENGTH: 1141
:
: TYPE: DNA
:
: ORGANISM: Artificial sequence
:
: FEATURE:
: NAME/KEY: promoter
: LOCATION: (1)..(1141)
:
: OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

```

Query Match	2.7%	Score 64.2	DB 4	Length 1141
Best Local Similarity	12.7%	Pred. No. 4.3e-07		
Matches 136	Conservative 379	Mismatches 550	Indels 10	Gaps 3
Qy	830	TTTTTCGATTATCTTCAACATTAAACCTGGAGATCTATGGAAATCAACACGTTGGG	889	
Db	1129	TTTKTKTKYKANNNNNNNNNGMKDNRDADTKMSATGTAAWMTNHAKBGATMCMYMYWTSTN	1070	
Qy	890	TAAATTCACCTTAGCAGACACATCACTACTGTAGAAATGGACAGAAACAGAGCAATTCCT	949	
Db	1069	RRCMCMTYARATWYTRSNANWSCATKXMMWTMKYATKYRATWYMAAMCWMNNNNWCCATN	1010	
Qy	950	GAATGGGCTATTAATATAGAGAAATACGTAAGAGTGTCCTGTAATTAAGTACTCTATTAAA	1009	
Db	1009	GYAKSCATNNAMMYATTTRMAAYAAKAPARAAGNNMRATGAAAGKMGCMAMATYGBWVA	950	
Qy	1010	GAGTGAGACACGAATGGAGAAATATCATCCCAATTTCTAGCTCAGACTAGACTCGAA	1065	
Db	949	DTAGKMCNNNNNNWTTEDVRBMAKAKKNNNNNNAATWTCYRPAATNNNKAATMMMKWTHGAH	890	
Qy	1070	GGTTTCGAAACGTAACCGAG-----TTTCCCAACTACCTGTGATGTTTCAGTGAT	1122	
Db	889	SKKRTIRHHTTCRTTKTKNNNNNNNARITYYWHHAARRWMAWMTTNNNNNNNNNACANT	830	
Qy	1123	CTTTCATCTCATGCTTATTAATGAGAGTAAGATGATTTCTCACCAAAATTGAATGACAA	1182	

[illegible]

```

US-09-806-708B-22

Query Match      2.4%; Score 57.8; DB 4; Length 1141;
Local Similarity 10.9%; Fred. No. 2.3e-05;
Matches 82; Conservative 263; Mismatches 407; Indels 0; Gaps 0;

OY 1565 TGTTCATTAATTCCTTTGGCAAAGCTGCTGACTTACAGAAAGTCTGCTTACAGACAGAA 1624
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 201 DRDMSBKRMVYGMGBMKWMSYDYVYVYVWDMDCRKXRYRWRRTGMRMKNYVAMBTAR 260
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1625 AGATGGCATTTTAAACGGTACTTGAAAATAAGAAAAATGAGAACTATGAAAAATTCATG 1684
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 261 RRYNNGMTBAAAYRMTWNNNNNNNAKMKCRKRYGMRBAYVNSTCTTWSKXTTKVTS 320
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1685 AAGCAATGGGTAAAGCTTACTTTTGAATGCTCTTAAAGACAGATACCACTACGGGG 1744
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 321 WANNCRADDAKDRHKMKWMSAMGVYNNNNNNNNNTYKKAIRBAMDWVHSAKMHAN 380
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1745 GAATACAAACTTAAGCTGTCTATGAACTACCACTCGGGCTAACCTGTCCTTGTGCTGC 1804
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 381 AAHSRKKWTEYKRTTWNNNNGTTTMMKMMAMMYKMDMDGTTNNNNNGRTTYGWRK 440
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1805 TATTTTGCCCTTGCACATTTGCCCTGCACCTTATTTTAAAAAGCTATAGAGGGAATAC 1864
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 441 NKQWVTTYKMKANNCKMRAMDHKCTHTNNTTMMKKKTYWNNCYWMSMTGKSHBBAAY 500
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1865 AAGGAGAAAAACATCTGATTTTATTTGCACTGGCATPATCTTATGCACTTTAGCTAAT 1924
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 501 TWYMMWRVYAHANNMNDYWKACTWYKBYVCSKMNMYAAMYKSSWNYTSRYVRMT 560
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1925 CCAGTAGGAGCATCCACAGAAATTTAAATGATTTATGTAAGAAATATTATTATTGA 1984
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 561 NNSWRKSDTSMGRANNYARABHYGKMTNRWMSHTBHRGAALYHMBMYBAK 620
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1985 TAAAGCTGTTGAAAAAATTTACACAAGAGGAAATTCGTGCTCCAGTTTTCAGAAACAC 2044
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 621 HOMKAWYAKKYAGAGSGSNNNNNNNNNNNNNNNNNATCARDDVYASRWYAMAAKMYRK 680
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 2045 ACATGATTTAGTCACTTTTAACTAGCTAGTCTTACTTAACTGTGACAACTGCTGT 2104
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 681 BAANNAYYTHANNWGCWNNNATDTRRTTWKNNNNNAGWKNNNNNNANNAKAAYKAAA 740
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 2105 AATATGATGTAAACATATCTACTAGTGTGATAGTAGTGTGATTAACGGCTGAAC 2164
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 741 AVKAKKRWANRKMAMRGWHDABAABTTDKNNNGAYTYTNNNNYRGVNTTAAD 800
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 2165 TGCCTCAGTGAAGAGGTGAGAAAGATAGACTGTGACAGATTTCTGGCTAAGCTCCC 2224
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 801 GMANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 860
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 2225 TCAACTACAAAAAGTCAACATATAAATGCAACATGATGTTCTATTTGTTTCTCT 2284
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 861 WRBAYTNNNNNNRMYAGAYADVAYVMSDTCDAWKKMWDATKNNATYTRGTARPTNN 920
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 2285 GCTTGATGTTAATGATTAATTAATTTT 2316
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 921 NNNMTKTKRYBHAAAMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 952
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-09-949-016-11761
; Sequence 11761, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```


EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 474
LENGTH: 515
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1) (515)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-474

Query Match 1.9%; Score 44.2; DB 4; Length 515;
Best Local Similarity 75.3%; Pred. No. 0.076;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1623 AAGATGCGATTAAACGCTTGAAGAAATGAGAAATGAGAACTATGAAATTCAT 1682
DB 23 AATCATGGCGTTGACAGCAGCTTGAAGCTAGACCGAGTGAATCTATGACAACTTCAT 82
QY 1683 GGAAGCAATGGGT 1695
DB 83 GGAATAAATGGGT 95

RESULT 7
US-09-949-016-4519
Sequence 4519 Application US/09949016
Patent No. 6812338
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 4519
LENGTH: 2252
TYPE: DNA
ORGANISM: Human
US-09-949-016-4519

Query Match 1.9%; Score 44.2; DB 4; Length 2252;
Best Local Similarity 75.3%; Pred. No. 0.16;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1623 AAGATGCGATTAAACGCTTGAAGAAATGAGAAATGAGAACTATGAAATTCAT 1682
DB 58 AATCATGGCGTTGACAGCAGCTTGAAGCTAGACCGAGTGAATCTATGACAACTTCAT 117
QY 1683 GGAAGCAATGGGT 1695
DB 118 GGAATAAATGGGT 130

RESULT 8
US-09-949-016-19
Sequence 19 Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 19
LENGTH: 2273
TYPE: DNA
ORGANISM: Human
US-09-949-016-19

Query Match 1.9%; Score 44.2; DB 4; Length 2273;
Best Local Similarity 75.3%; Pred. No. 0.16;
Matches 55; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1623 AAGATGCGATTAAACGCTTGAAGAAATGAGAAATGAGAACTATGAAATTCAT 1682
DB 58 AATCATGGCGTTGACAGCAGCTTGAAGCTAGACCGAGTGAATCTATGACAACTTCAT 117
QY 1683 GGAAGCAATGGGT 1695
DB 118 GGAATAAATGGGT 130

RESULT 9
US-08-232-463-14/c
Sequence 14 Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOXP2 VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMNU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F15
US-08-232-463-14

Query Match
Best Local Similarity 1.8%; Score 43.8; DB 1; Length 7218;
Matches 36; Conservative 204; Mismatches 191; Indels 0; Gaps 0;

QY 628 TATCTGTGATATAGCAACATTTCACTTACAGATTACAGATTACATTTTGAAGCTATA 687
   |||||
Db 1474 TATCTGTGATATAGCAAGATTAAGATTAAGATTAAGATTTGTACRRRRRRRRRRRRR 1415

QY 688 GACAGACAGATTGGTGGCCGTCATAGACAGACTAATCTAGATTAGTACAA 747
   :::::
Db 1414 GACAGATTTAGACAGCAATTTGCTCAGTTCAAGTACATATCTTGGGGAGAGAG 807
   :::::
QY 748 GACAGATTTAGACAGCAATTTGCTCAGTTCAAGTACATATCTTGGGGAGAGAG 807
   :::::
Db 1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295

QY 808 GACCTGAGCCAGTGTGCTCATTTTCTGCATTATCTTCAACATTTAAACCTGGGATC 867
   :::::
Db 1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235

QY 868 TATGGAATCAACACAGCTGGTAAATTCATCTAGACAGACATCACTAGTGAAGAT 927
   :::::
Db 1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175

QY 928 GACAGACAGACATTTCACTGATGGCTATATATAGAAATCGTGAAGCTGCC 987
   :::::
Db 1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

QY 988 TGAATTAGCTACTATTAAAGATGAGACAGACAGATGAGATATCATCGCATTTCT 1047
   :::::
Db 1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1055

QY 1048 GTAGCTCAGCA 1058
   |||||
Db 1054 CTCGACTGCA 1044

RESULT 10
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 1.8%; Score 43; DB 4; Length 832;
Matches 51; Conservative 147; Mismatches 142; Indels 1; Gaps 1;
```

```
QY 1073 TTCTGAACCTGACACGAGTTTCCCAAACTACCTGTGAGTTCAGTGCATCTCATCTC 1132
   :::::
Db 35 KTYRWRRRKKKKKMMWTKWTWYRPAWGGYKKKAKMCRITKTKKKKKGGYMMWYMGW 94

QY 1133 ATGCTTATTATGTGAGATAGATATCTCACCAATTAGATGACAAAGAGAGATT 1192
   :::::
Db 95 RRSYMAWMTRTWGTAYVYSMMYRWYRCMKKAYRKTCYSSKSGMTWKKRKKMTTW 154

QY 1193 TGTGTTTT-ATCTGTGGGTAATACGTTTCTCCAGTTGTATTAAGACCCCTCCACAG 1251
   :::::
Db 155 WMKKTYWMAATRYMMWMTKWRASWYCMWNGKARKSTYKSRSSYASRSKRCYCS 214

QY 1252 TATAAGTCTATGCAACAAAGAAATGCAATACATCTCTAGTCTCATTTATTTT 1311
   :::::
Db 215 CSWGAWSWKYMMWMMRWGATGAGMKAMPDASCMBRKAYAGSKTSYKSMWMTSRWKY 274

QY 1312 CATTAATAGCCGGTTTTTTTACTACACATCAATTAAGATGACAGATGAGGATAGT 1371
   :::::
Db 275 CYTKAWTGYCYCRKGGMMCKRGRTYASKYMMKRWMCWARYRSTGTASMMWRWY 334

QY 1372 GACTGTTATTAAGAAGAGTAATTAAGATATCATCATTT 1412
   :::::
Db 335 YTWMMKMMWYKVAABAARWMMWMMWBRBACAATAATTAATT 375

RESULT 11
US-09-949-016-18019/c
; Sequence 18019, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FaestSeq for Windows Version 4.0
; SEQ ID NO 18019
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18019

Query Match
Best Local Similarity 1.8%; Score 42.6; DB 4; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 2218 AGCTCCCTCACTACGAAAAAGTCACATTAATAAAGCAACATGATGTTTGT 2277
   |||||
Db 258 AGTTTCTTCCCGCCCTCAATAAACAACAACATCACTGAAAAATAAACACTTC 199

QY 2278 TTCTCTGCTGATGTATATGATATTTATTTATTTTATTTAGCGGTAAGTATGA 2337
   |||||
Db 198 CTATGGGATTTGACTTATTTTCTCAATGCTTACCTTTACAGGTGTTAATATAGTGA 139

QY 2338 AAAGAAAGTTAGAGAGCCACGATTAATCTGAAGCTCACTATT 2378
   |||||
Db 138 AAAGAAAGCTGACAGCTCATGACAAATTTGAAGCTGACAAATT 98

RESULT 12
US-09-949-016-18020/c
; Sequence 18020, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 18020
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18020

Query Match
Best Local Similarity 54.0%; Score 42.6; DB 4; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```

QY 2218 ACCTCCCTCACTACAGAAAAGTCAATAAATGCAACATGATGTTCTATTGTT 2277
DB 554 AGTTTCTTCCACCTCAAAATTAACAACAAACATTAAGTGAACAACTTC 495
QY 2278 TTTCCTGCTGATGTAATGATTAATTAATTTTATTTTATAGCGGGAATGATGA 2337
DB 494 CTATGGATTGACTTATTTCTCCATTGCTTACCTTTACAGGTGTTAATATAGTGA 435
QY 2338 AAGAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 434 AAGGAGCTTGACGCTCATGACAAATTGGAAGCTGACATT 394
```

```

RESULT 13
US-09-949-016-161279/c
; Sequence 161279, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 161279
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-161279
```

```

Query Match
Best Local Similarity 54.0%; Score 42.6; DB 4; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 2218 ACCTCCCTCACTACAGAAAAGTCAATAAATGCAACATGATGTTCTATTGTT 2277
DB 258 AGTTTCTTCCACCTCAAAATTAACAACAAACATTAAGTGAACAACTTC 199
QY 2278 TTTCCTGCTGATGTAATGATTAATTAATTTTATAGCGGGAATGATGA 2337
DB 198 CTATGGATTGACTTATTTCTCCATTGCTTACCTTTACAGGTGTTAATATAGTGA 139
```

```

QY 2338 AAGAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 139 AAGGAGCTTGACGCTCATGACAAATTGGAAGCTGACATT 98
```

```

RESULT 14
US-09-949-016-161280/c
; Sequence 161280, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 161280
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-161280
```

```

Query Match
Best Local Similarity 54.0%; Score 42.6; DB 4; Length 601;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 2218 ACCTCCCTCACTACAGAAAAGTCAATAAATGCAACATGATGTTCTATTGTT 2277
DB 554 AGTTTCTTCCACCTCAAAATTAACAACAAACATTAAGTGAACAACTTC 495
QY 2278 TTTCCTGCTGATGTAATGATTAATTAATTTTATAGCGGGAATGATGA 2337
DB 494 CTATGGATTGACTTATTTCTCCATTGCTTACCTTTACAGGTGTTAATATAGTGA 435
QY 2338 AAGAAGTTAGAGCCCAAGATTAATCTGAAGCTCACTATT 2378
DB 434 AAGGAGCTTGACGCTCATGACAAATTGGAAGCTGACATT 394
```

```

RESULT 15
US-09-385-982-376/c
; Sequence 376, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 376
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(611)
```


OTHER INFORMATION: n = A,T,C or G
US-09-385-982-376

Query Match 1.8%; Score 42.6; DB 3; Length 611;
Best Local Similarity 55.0%; Pred. No. 0.23;
Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 2220 CTCCTCACTAGAAAAAGTCACATATAAATGCAACATGATGTTCTATTGTTT 2279
Db 351 CACNTTATATTCNGTTTACCGAAGAAAAAGCGGACNTTTTTTTTTTTTTT 292
QY 2280 TCTCTGCTGATGTTAATGATTATTTATTTTTTTTAGGCGGAATGTGATAA 2339
Db 291 TTTTTTTTGNTTTTTTTTTTTTTTTTTTTTTTTATAAGATTATTAGATATAA 232
QY 2340 AGAAGTTAGG 2350
Db 231 GGGGAGATTAGG 221

Search completed: November 21, 2005, 02:06:35
Job time : 419 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2005, 00:03:25 ; Search time 1699 Seconds

(without alignments)
11589.429 Million cell updates/sec

Title: US-10-099-663-1

Perfect score: 2381

Sequence: 1 agctctctgcgcgaagaag.....atctgaagctcactatcag 2381

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9794790 seqs, 413490567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the total being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2381	100.0	2381	17	US-10-099-663-1	Sequence 1, Appli
2	336	14.1	336	17	US-10-099-663-2	Sequence 2, Appli
3	50.6	2.1	16914	20	US-10-741-601-5698	Sequence 5698, Ap
4	50.6	2.1	16914	22	US-10-741-600-17777	Sequence 17777, A
5	49.8	2.1	630	14	US-10-027-632-5694	Sequence 5694, Ap

C	6	49.8	2.1	630	18	US-10-027-632-5694	Sequence 5694, Ap
C	7	49.8	2.1	5204	18	US-10-116-275-303	Sequence 303, App
C	8	48	2.0	546	13	US-09-925-065A-51803	Sequence 51803, A
C	9	48	2.0	569	13	US-09-925-065A-200196	Sequence 200196, A
C	10	48	2.0	570	13	US-09-925-065A-200197	Sequence 200197, A
C	11	48	2.0	814	13	US-09-925-065A-63272	Sequence 63272, A
C	12	47.2	2.0	337	9	US-09-925-065A-63272	Sequence 63272, A
C	13	47	2.0	2636	26	US-11-097-145-21061	Sequence 21061, A
C	14	46.8	2.0	662	21	US-10-335-053-36	Sequence 36, Appl
C	15	46.8	2.0	670	22	US-10-764-420-2245	Sequence 2245, Ap
C	16	46.2	1.9	6222	16	US-10-311-455-665	Sequence 665, App
C	17	44.4	1.9	565	13	US-09-925-065A-110075	Sequence 110075, A
C	18	44.2	1.9	515	10	US-09-871-161-474	Sequence 474, App
C	19	44.2	1.9	2252	20	US-10-741-601-261	Sequence 261, App
C	20	44.2	1.9	2252	22	US-10-741-600-745	Sequence 745, App
C	21	44.2	1.9	3673778	17	US-10-312-841-2	Sequence 2, Appli
C	22	43.6	1.8	400	21	US-10-425-115-11538	Sequence 11538, A
C	23	43.6	1.8	433	20	US-10-674-124A-3159	Sequence 2159, Ap
C	24	43.6	1.8	14023	16	US-10-311-455-2077	Sequence 2077, Ap
C	25	43.2	1.8	7143	16	US-10-311-455-956	Sequence 956, App
C	26	42.8	1.8	558	13	US-09-925-065A-572321	Sequence 572321, A
C	27	42.8	1.8	679	13	US-09-925-065A-875384	Sequence 875384, A
C	28	42.8	1.8	679	13	US-09-925-065A-875385	Sequence 875385, A
C	29	42.8	1.8	679	13	US-09-925-065A-875386	Sequence 875386, A
C	30	42.8	1.8	51289	20	US-10-322-281-648	Sequence 648, App
C	31	42.6	1.8	513	17	US-10-029-386-13006	Sequence 13006, A
C	32	42.6	1.8	611	10	US-09-871-161-376	Sequence 376, App
C	33	42.6	1.8	1241	21	US-10-425-115-124424	Sequence 124424, A
C	34	42.6	1.8	2844	15	US-10-176-847-19	Sequence 19, Appl
C	35	42.6	1.8	5501	16	US-09-311-455-28	Sequence 28, Appl
C	36	42.4	1.8	452	9	US-09-922-217-138	Sequence 138, App
C	37	42.4	1.8	452	9	US-09-833-263-138	Sequence 138, App
C	38	42.4	1.8	452	14	US-10-025-380-138	Sequence 804505, A
C	39	42.4	1.8	479	13	US-09-925-065A-804505	Sequence 804505, A
C	40	42.4	1.8	481	13	US-09-925-065A-806667	Sequence 806667, A
C	41	42.4	1.8	641	13	US-09-925-065A-804504	Sequence 804504, A
C	42	42.4	1.8	682	15	US-10-198-848-1721	Sequence 1721, Ap
C	43	42.4	1.8	5908	15	US-10-239-676-94	Sequence 94, Appl
C	44	42.4	1.8	5908	16	US-10-240-453-106	Sequence 106, App
C	45	42.4	1.8	5908	18	US-10-221-613-176	Sequence 176, App

ALIGNMENTS

RESULT 1
US-10-099-663-1
Sequence 1, Application US/10099663
Publication No. US20030177516A1
GENERAL INFORMATION:
APPLICANT: Avigenics, Inc
TITLE OF INVENTION: Avian Gut-Specific Promoters
FILE REFERENCE: A181
CURRENT APPLICATION NUMBER: US/10/099,663
CURRENT FILING DATE: 2002-03-14
NUMBER OF SEQ ID NOS: 19
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 2381
TYPE: DNA
ORGANISM: Gallus gallus
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)..(1626)
NAME/KEY: exon
LOCATION: (1627)..(1693)
NAME/KEY: Intron
LOCATION: (1694)..(2322)
NAME/KEY: exon
LOCATION: (2323)..(2381)
US-10-099-663-1
Query Match 100.0%, Score 2381, DB 17, Length 2381.

QY	2161	ACACTGCTCAGTGAAGAAGTGAAAGCTGCAATCGAATTCTGGGCTAAC	2220
Db	2161	ACACTGCTCAGTGAAGAAGTGAAAGCTGCAATCTGGGCTAAC	2220
QY	2221	TCCCTCAACTACAAAAAGTCACAAATCAAATGTTCTATTTGTGTTTT	2280
Db	2221	TCCCTCAACTACAAAAAGTCACAAATCAAATGTTCTATTTGTGTTTT	2280
QY	2281	CTCGCTTGATGTTAATTGATTATTTTCTTTTAAAGCGGTGAATGTATGAAA	2340
Db	2281	CTCGCTTGATGTTAATTGATTATTTTCTTTTAAAGCGGTGAATGTATGAAA	2340
QY	2341	GAAAGTAGAGCCACGATAATCGAAGCTCACGATTTCAAG	2381
Db	2341	GAAAGTAGAGCCACGATAATCTGAAGCTCACGATTTCAAG	2381

```

RESULT 2
US-10-099-663-2
; Sequence 2, Application US/10099663
; Publication No. US20030177516A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics, Inc
; TITLE OF INVENTION: Avian GUT-Specific Promoters
; FILE REFERENCE: A181
; CURRENT APPLICATION NUMBER: US/10/099,663
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(336)
; OTHER INFORMATION: chick intestinal fatty acid binding protein promoter region
; US-10-099-663-2

```

Query Match	14.1%	Score 336	DB 17	Length 336
Best Local Similarity	100.0%	Pred. No. 5.4e-74		
Matches 336	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1301	ATTATTATTTTCATTAGATAGCCGGTTTTTTACTACAACTCAATTAAGATGAACAGATG	1360	
Db	1	ATTATTATTTTCATTAGATAGCCGGTTTTTTACTACAACTCAATTAAGATGAACAGATG	60	
QY	1361	AATGGGTTAGTACTGCTGTTTATAAAGAAAGTAATTAAGATCTATCATCATTTGAGCGAA	1420	
Db	61	AATGGGTTAGTACTGCTGTTTATAAAGAAAGTAATTAAGATCTATCATCATTTGAGCGAA	120	
QY	1421	TAAAGGAGGAGAGAGATTCAGCAAAACAATGTCCTTACAAGTGGAAAAACAATTAAACTTAAA	1480	
Db	121	TAAAGGAGGAGAGAGATTCAGCAAAACAATGTCCTTACAAGTGGAAAAACAATTAAACTTAAA	180	
QY	1481	GTCAGCCCCCTCCTCTTGCACAAGATCAATGSCCACAAGTTAGCTTTGAGCCAGCCACATCATCA	1540	
Db	181	GTCAGCCCCCTCCTCTTGCACAAGATCAATGSCCACAAGTTAGCTTTGAGCCAGCCACATCATCA	240	
QY	1541	TGTAAATTTGCTTTCCTGATTAAGCCTGTTTCATTAATTTCTTTTGCAAAAGCTCTGCTACTTA	1600	
Db	241	TGTAAATTTGCTTTCCTGATTAAGCCTGTTTCATTAATTTCTTTTGCAAAAGCTCTGCTACTTA	300	
QY	1601	CCAGAACTCTGCCCTACAGACAGAAAAGATGGCATTTTA	1636	
Db	301	CCAGAACTCTGCCCTACAGACAGAAAAGATGGCATTTTA	336	

RESULT 3
US-10-741-601-5698
; Sequence 5698, Application US/10741601
; Publication No. US20040166519A1

	GENERAL INFORMATION:	Michele et al.	
	APPLICANT: CARGILL,		
	TITLE OF INVENTION:	GENETIC POLYMORPHISMS ASSOCIATED WITH	
	TITLE OF INVENTION:	STENOSIS, METHODS OF DETECTION AND USES THEREOF	
	FILE REFERENCE:	CL001500	
	CURRENT APPLICATION NUMBER:	US/10/741,601	
	CURRENT FILING DATE:	2003-12-22	
	NUMBER OF SEQ. ID NOS:	26415	
	SOFTWARE:	FastSeq for Windows Version 4.0	
	SEQ ID NO 5698		
	LENGTH:	16914	
	TYPE:	DNA	
	ORGANISM:	Homo sapiens	
	US-10-741-601-5698		
QY	Query Match	2.1%; Score 50.6; DB 20; Length 16914;	
	Best Local Similarity	58.8%; Pred. No. 0.46;	
	Matches 124; Conservative	2; Mismatches 76; Indels 9; Gaps 2;	
DB	1516 TGAAGCTTTGACCAAGCCATCATCATCATTAATTGCTTTCCTGATAAGCCTGTTCATTAAT	1575	
	5945 TGAACCTTTAAGCTTCCACATCACATATATGAAGTTGGTTAAAGATAAATAATATATTAAT	6004	
OY	1576 TCAC-----TTTGCAAAGCTCTGCTACTTTACCCAGAAGCTCGCTACAGACAGANA	1627	
	6005 TCTTCGCCCAAGACAGACAGACTGAATCTTACTCTGCTTAGAGGCTGACT-CAACTGAAATCA	6063	
DB	1628 TGGCATTTAAACGCTACTTTGAAAATAGAAAAATAGAACTATGAAAAATTCATGAG	1687	
	6064 TTGCGTTTGACACACACTTGGAAAGTAGACCGAGATGAAAACTATGACAAGTTATGGA	6123	
OY	1688 CAATGGGTAAAGCCTTACTTTTTTGAATGCTT	1718	
	6124 AATGGGTAAAGACTTATTTCTTTGGGCT	6154	

```

RESULT 4
US-10-741-600-17777
; Sequence 17777, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CI001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ. ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 17777
; LENGTH: 16914
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17777

Query Match          2.1%; Score 50.6; DB 22; Length 16914;
Best Local Similarity 58.8%; Pred. No. 0.46;
Matches 124; Conservative 2; Mismatches 76; Indels 9; Gaps 2;

```

[illegible]

Db 6124 AATGGGTAAAGACTTATTCTTTGTGGCT 6154

RESULT 5

US-10-027-632-5694/c
; Sequence 5694, Application US/10027632
; Publication No. US20020198371A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129

CURRENT FILING DATE: 2002-04-30

PRIOR FILING DATE: 2000-07-12

PRIOR FILING DATE: 2000-07-12

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-03-29

PRIOR FILING DATE: 2000-03-29

PRIOR FILING DATE: 2000-02-24

PRIOR FILING DATE: 1999-11-23

PRIOR FILING DATE: 1999-09-28

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5694

LENGTH: 630

TYPE: DNA

ORGANISM: Human

US-10-027-632-5694

Query Match 2.1%; Score 49.8; DB 14; Length 630;

Best Local Similarity 59.2%; Pred. No. 0.12; Indels 9; Gaps 2;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

Db 1516 TGAGCTTTAGCCAGCATCATGTAATGCTTCTGTAAGCTGTTCAAT 1575

Db 531 TGAACCTTAAGCTTCCACATCAGATGTAAGTTGTTCAAGATAAGAAATATATTAAT 472

QY 1576 TCTC-----TTGCAAGCTCTGCTACTTACCAAGTCTGCTTACGAGAGAA 1627

Db 471 TCTGCCCAAGAGCAGACCTGAACTCTTACCTGCTTACGAGCTGACT-CAACTGAATCA 413

QY 1628 TGACATTTAAGCTTACCTTGAATAATGAGAAATGAGAACTATGAAATTCATGGAAG 1687

Db 412 TGGCGTTTGACAGCACTTGAAGGTAGACCGAGTGAACCTATGACAACTTATGAGAA 353

QY 1688 CAATGGGTAAAGCTTACTTTTGAATGCT 1718

Db 352 AATGGGTAAAGACTTATTCTTTGTGGCT 322

RESULT 6

US-10-027-632-5694/c
; Sequence 5694, Application US/10027632
; Publication No. US20030204075A9

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE REFERENCE: 108827.129

CURRENT FILING DATE: 2002-04-30

PRIOR FILING DATE: 2000-07-12

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-04-20

PRIOR FILING DATE: 2000-03-29

PRIOR FILING DATE: 2000-02-24

PRIOR FILING DATE: 1999-11-23

PRIOR FILING DATE: 1999-09-28

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5694

LENGTH: 630

TYPE: DNA

ORGANISM: Human

US-10-027-632-5694

Query Match 2.1%; Score 49.8; DB 18; Length 630;

Best Local Similarity 59.2%; Pred. No. 0.12; Indels 9; Gaps 2;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCATCATGTAATGCTTCTGTAAGCTGTTCAAT 1575

Db 531 TGAACCTTAAGCTTCCACATCAGATGTAAGTTGTTCAAGATAAGAAATATATTAAT 472

QY 1576 TCTC-----TTGCAAGCTCTGCTACTTACCAAGTCTGCTTACGAGAGAA 1627

Db 471 TCTGCCCAAGAGCAGACCTGAACTCTTACCTGCTTACGAGCTGACT-CAACTGAATCA 413

QY 1628 TGACATTTAAGCTTACCTTGAATAATGAGAAATGAGAACTATGAAATTCATGGAAG 1687

Db 412 TGGCGTTTGACAGCACTTGAAGGTAGACCGAGTGAACCTATGACAACTTATGAGAA 353

QY 1688 CAATGGGTAAAGCTTACTTTTGAATGCT 1718

Db 352 AATGGGTAAAGACTTATTCTTTGTGGCT 322

RESULT 7

US-10-116-275-303

; Sequence 303, Application US/10116275

; Publication No. US20030211476A1

GENERAL INFORMATION:

APPLICANT: Egan Pharmaceutical Technology

APPLICANT: O'Mahony, Daniel J.

APPLICANT: Brayden, David

APPLICANT: Byrnie, Dargagh

APPLICANT: Lambkin, Imelda

TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and

FILE REFERENCE: E1067/20087

CURRENT FILING DATE: 2002-10-04

NUMBER OF SEQ ID NOS: 349

SOFTWARE: Patentin version 3.1

SEQ ID NO 303

LENGTH: 5204

TYPE: DNA

ORGANISM: Homo sapiens

US-10-116-275-303

Query Match 2.1%; Score 49.8; DB 18; Length 5204;

Best Local Similarity 59.2%; Pred. No. 0.38; Indels 9; Gaps 2;

Matches 125; Conservative 0; Mismatches 77; Indels 9; Gaps 2;

QY 1516 TGAGCTTTAGCCAGCATCATGTAATGCTTCTGTAAGCTGTTCAAT 1575

Db 971 TGAACCTTAAGCTTCCACATCAGATGTAAGTTGTTCAAGATAAGAAATATATTAAT 1030

QY 1576 TCTC-----TTGCAAGCTCTGCTACTTACCAAGTCTGCTTACGAGAGAA 1627

Db 1031 TCTGCCCAAGAGCAGACCTGAACTCTTACCTGCTTACGAGCTGACT-CAACTGAATCA 1089

; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.135
 ; CURRENT ADDITION NUMBER: 115/00/0035 0053

FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A

```

; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63272
; LENGTH: 814
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-63272
```

```

Query Match
Best Local Similarity 2.0%; Score 48; DB 13; Length 814;
Matches 66; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```

```

QY 1623 AAGATGCGATTAAAGCTTCTGAAATAGAGAAATGAGAACTATGAAATTCAT 1682
    |||||
DB 415 AATCATGGCGTTTGACAGACCTTGGAGGTAGACCGAGTGAATATGACAGTTTCA 356
    |||||
QY 1683 GGAAGCAATGCGTACCTTCTTTTGAATGCTT 1718
    |||||
DB 355 GGAATAATGGGTAAAGACTTATTTCTTGTGGCT 320
    |||||
```

```

RESULT 12
US-09-960-352-6036/c
; Sequence 6036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6036
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3058-015-Q1-K1-G5
US-09-960-352-6036
```

```

Query Match
Best Local Similarity 2.0%; Score 47.2; DB 9; Length 337;
Matches 97; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
```

```

QY 2185 AAGATGAGCTGTGAGTCAATCTGGCTAAGCTCCCTCACTACAGAAAAAGTCA 2244
    |||||
DB 289 AATTAATATTAATATTAATTAACAATTAATTAAGTTACTATTAATAAAAATGTGAT 230
    |||||
QY 2245 AATAAATGCAACATGATGCTATTTGTTTCTGCTGATGATGATGATGAT 2304
    |||||
DB 229 AATTAATATTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTT 170
    |||||
QY 2305 TATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2364
    |||||
DB 169 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 110
    |||||
```

```

RESULT 13
US-11-097-143-21061
```

```

; Sequence 21061, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
```

```

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: C1000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21061
; LENGTH: 2636
; TYPE: DNA
; ORGANISM: DROSOPHILA
```

```

US-11-097-143-21061
```

```

Query Match
Best Local Similarity 2.0%; Score 47; DB 26; Length 2636;
Matches 104; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
```

```

QY 1821 ATGCGCTGCACTTATTTTGAAGAAAGCTTATAGAGGGGAATACAGAGAGAAAAACAT 1880
    |||||
DB 540 ATTGACTTGAATATCTTAAATATACAGCCCTGTTGCTTCTTGAATTAACAT 599
    |||||
QY 1881 CTGATTTATTTGCAATGCGATATCTTATGCAATTAAGTATTCAGTAGAGCATTC 1940
    |||||
DB 600 ATTAATTTATTTAAATGCGAATTAATTTGCTTAAATTAATTAATTTATTTTTC 659
    |||||
QY 1941 AGCAGAAATTTAATAGATTATATATAGAAATATTTATTTGATTAAGCTGTTGAAA 2000
    |||||
DB 660 TTAGGGAATATAAGATTATTTGCTGCTATGTTTATATTTAATTAATTAATTAAGCA 719
    |||||
QY 2001 ATTACACAAGAGGAAT 2019
    |||||
DB 720 ATTAGTAGATGTGTAATTT 738
    |||||
```

```

RESULT 14
US-10-335-053-36
; Sequence 36, Application US/10335053
; Publication No. US20040241653A1
; GENERAL INFORMATION:
; APPLICANT: Quark Biotech, Inc.
; TITLE OF INVENTION: Method for identifying marker genes for cancer
; FILE REFERENCE: 68733-A, 070/US1
; CURRENT APPLICATION NUMBER: US/10/335,053
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: 60/345,317
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 36
; LENGTH: 662
; TYPE: DNA
```


100% (uspto)